Check for updates

Development and Validation of KN-DIOC: A Novel Preoperative Diagnostic Index Using Ultrasound, Complete Blood Count, and Cancer Antigen 125 for Ovarian Cancer

Sorawit Tongyib[a], Teerapol Saleewong[a, c], Woraphot Chaowawanit[b]

Abstract

Background: Ovarian cancer, particularly epithelial ovarian cancer (EOC), is one of the deadliest gynecological malignancies due to nonspecific early symptoms and late diagnosis. Current diagnostic tools, while useful, often do not account for regional variations in disease presentation, particularly in Asian populations. This study aimed to develop and validate a new preoperative diagnostic index tailored to the Thai population by integrating complete blood count (CBC), tumor markers, and ultrasound features.

Methods: This retrospective cohort study included patients with pathologic pelvic or adnexal masses scheduled for surgery at Vajira Hospital from April 2022 to October 2024. Clinical data, CBC, cancer antigen 125 (CA125) levels, and ultrasound findings were analyzed to develop and validate a diagnostic index (KMUTT-NMU Diagnostic Index for Ovarian Cancer (KN-DIOC)). The model's performance was compared against established indices like Risk of Malignancy Index (RMI), Risk of Ovarian Malignancy Algorithm (ROMA), and Rajavithi-Ovarian Cancer Predictive Score (R-OPS) through multivariate logistic regression, focusing on key predictors.

Results: The study comprised 195 patients divided into 151 for the development dataset and 44 for the validation dataset. The KN-DIOC showed high discriminative ability with an area under curve (AUC) of 0.866, indicating very good capability in differentiating between benign and malignant ovarian masses. The index achieved a sensitivity of 93.75% and a specificity of 78.57%, demonstrating superior performance to traditional diagnostic tools, especially in the validation dataset.

Conclusion: The novel diagnostic index (KN-DIOC), incorporating CBC, ultrasound features, and tumor markers, provides a robust tool for preoperative assessment of ovarian tumors in Thai patients. It offers significant improvements in sensitivity and specificity over existing models, suggesting its potential for broader application in similar settings. This index supports enhanced decision-making in gynecological oncology, potentially leading to better patient outcomes through timely and accurate diagnosis.

Keywords: Ovarian cancer; Predictive model; Complete blood count; Tumor marker; Ultrasound

[a]Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand
[b]Department of Obstetrics and Gynecology, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok 10300, Thailand
[c]Corresponding Author: Teerapol Saleewong, Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi, , Bangkok 10140, Thailand. Email: teerapol.sal@kmutt.ac.th

Introduction

Ovarian cancer, particularly epithelial ovarian cancer (EOC), ranks as the eighth most prevalent cancer in women worldwide, recording 313,959 new cases and resulting in 207,252 fatalities in the year 2020 [1]. Ovarian cancer is notorious for its vague and nonspecific early symptoms, often leading to delayed diagnosis and contributing significantly to its high fatality rate [1]. The challenge in early detection contributes to ovarian cancer being the leading cause of death from gynecological malignancies. Diagnosis typically involves comprehensive medical history assessment, physical examination, and transvaginal ultrasound, the latter of which, while critical for initial assessments, cannot conclusively determine malignancy. Instead, it offers vital diagnostic clues [2].

In 1990 [2], Jacobs et al described specific morphological patterns or ultrasound features associated with ovarian cancer, including multilocular cysts, solid areas, bilateral lesions, ascites, and metastasis. Based on serum cancer antigen 125 (CA125) levels, menopausal status, and ultrasound features, the Risk of Malignancy Index (RMI) was developed [2]. These features must be assessed by an experienced ultrasound specialist, as even the slightest error can lead to a missed diagnosis and potentially serious medical consequences.

Differences in the skill level of operators and the interpretation of morphological characteristics of ovarian tumors can impact the accuracy of scoring systems. Moore and associates recognized human epididymis protein 4 (HE4) as the most sensitive biomarker for ovarian cancer and created the Risk of Ovarian Malignancy Algorithm (ROMA), which integrates HE4 and CA125 for risk classification [3]. The advantage of employing a

Table 1. Numbers of Histological Types in Development and Validation Datasets

Histological type	Development dataset (n = 53)	Validation dataset (n = 17)
High-grade serous carcinoma	19	4
Mucinous adenocarcinoma/carcinoma	8	1
Clear cell carcinoma	7	5
Endometrioid adenocarcinoma	8	5
Undifferentiated adenocarcinoma	7	0
Mixed cell type	3	1
Mucinous borderline tumor	1	1

serum biomarker algorithm is its objectivity, which removes the necessity for subjective assessments and thereby improves the consistency of results across various centers and regions.

Emerging from the efforts to improve diagnostic accuracy, the International Ovarian Tumor Analysis (IOTA) group developed the "10 simple rules" in 2008. This system classifies features into those indicative of malignant tumors (M-features) and those suggestive of benign tumors (B-features), aiding in the differentiation of ovarian masses [4].

The popular preoperative diagnostic indices include the Risk of Malignancy Index I-IV (RMI I-IV) [2, 5-7] and the ROMA [3]. These indices were developed to aid in preoperative diagnosis. However, the American College of Obstetricians and Gynecologists (ACOG) specifically recommends the RMI I as a referral tool for directing patients to gynecologic oncologists [8]. In Thailand, RMI is the preferred method for diagnosing ovarian cancer. However, a study [9] reported that RMI I and II have lower sensitivity in Thai patients. In addition, Rajavithi Hospital introduced the Rajavithi-Ovarian Cancer Predictive Score (R-OPS) in 2016, a novel scoring system developed using clinical data from Thai patients with pelvic masses. R-OPS incorporates menopausal status, serum CA125, serum HE4 levels, and ultrasound findings of solid lesions. The study demonstrated that R-OPS has higher diagnostic performance compared to RMI I and ROMA [10].

Differences in ethnicity, geography, diet, and other variables that may affect the features of ovarian cancer between the development sample and the Thai population may be the cause of this decrease. Additionally, compared to their European and American counterparts, Asian women exhibited distinct pathological distributions of ovarian cancers, with a higher prevalence of mucinous tumors, endometrioid carcinomas, and clear cell carcinomas (CCCs) [10].

The purpose of this study was to develop and validate a preoperative diagnostic index for Thai patients with EOC. To create a more useful diagnostic index for Thai people and people from comparable ethnic backgrounds, the index will include clinical data, tumor markers, and IOTA ultrasound features.

Materials and Methods

Study design and participant selection

The diagnostic prediction research, utilizing a retrospective

cohort, was conducted at Vajira Hospital, a center of excellence in cancer care in Thailand. The study protocol was approved by the Institutional Ethics Committee of the Faculty of Medicine, Vajira Hospital, as well as the Institutional Review Board (IRB) of KMUTT and Vajira Hospital. The approval document number is KMUTT-IRB-2022/0506/190.

All patients who presented with a pathologic pelvic or adnexal mass and were scheduled for surgery during April 2022 to October 2024 were enrolled. The operations were performed by experienced gynecologists, gynecologic oncologists, or residency/fellowship trainees under supervision. Patients who underwent surgery from April 2022 to May 2024 formed the "development dataset," which was used to develop a preoperative diagnostic index. The "validation dataset," which was used to validate the KMUTT-NMU Diagnostic Index for Ovarian Cancer (KN-DIOC), consisted of patients who underwent surgery from June to October 2024. Additionally, the diagnostic performance of our diagnostic index will be compared with RMI, ROMA, and R-OPS.

The selection criteria included patients diagnosed with benign ovarian tumors or EOC based on postoperative tissue pathology reports. Patients were categorized into two groups according to their diagnoses: EOC (case group) and benign ovarian tumors (control group). Borderline ovarian tumors were included in the benign group for patients confirmed to be non-malignant. The benign ovarian tumors considered in the study included dermoid cysts, endometriotic cysts, corpus luteal cysts, adenofibromas, adenomyomas, ovarian fibromas, paratubal cysts, leiomyomas, and hemorrhagic cysts. Patients diagnosed with metastatic cancer, non-EOC, recurrent ovarian cancer, or incomplete data were excluded. Within the ovarian cancer group, EOC was classified into different histological types, including high-grade serous carcinoma (HGSC), mucinous adenocarcinoma, CCC, endometrioid adenocarcinoma, undifferentiated adenocarcinoma, mixed cell type, and mucinous borderline tumor, across both datasets, as shown in Table 1.

Data collection

All clinical characteristics and potential predictors for diagnosis were collected from medical records. The data included information on menopausal status, age, blood test results (complete blood count (CBC), liver function tests (LFTs), etc.),

ultrasound features obtained by transvaginal ultrasound, tumor or cyst size, and levels of CA125 in U/mL. The ultrasound features included the presence of solid area, papillary projections, multilocular cyst, ascites, and the type of cyst wall, along with other features. The platelet-to-lymphocyte ratio (PLR) and lymphocyte-to-monocyte ratio (LMR) were additionally calculated from the collected data as the ratios of platelets to lymphocytes and lymphocytes to monocytes, respectively.

Postmenopausal women are defined as those who have experienced no menstrual period for more than 12 consecutive months (or 1 year) without any underlying pathological causes, or those older than 55 years.

RMI I-III was calculated using the product of the serum CA125 level (U/mL), ultrasound score (U), and menopausal status (M). These indices were defined as RMI I-III = U × M × CA125. RMI IV is calculated by adding the tumor size score (S) to the risk of malignancy indices, which is defined as RMI IV = U × M × CA125 × S, where S = 1 when the diameter is < 7 cm, and S = 2 when the diameter is ≥ 7 cm. The code and cutoff point for RMI I-IV and the ultrasound score were calculated as outlined in Supplementary Material 1 (wjon.elmerpub.com).

The ROMA was developed using serum levels of CA125 (U/mL), HE4 (pM/L), and menopausal status. The ROMA score was calculated in percentage using the formula: ROMA (%) = 100 × ($e^{PI}/(1 + e^{PI})$). The predictive index (PI) was defined separately based on menopausal status: for premenopausal, PI = -12 + 2.38 × ln(HE4) + 0.0626 × ln(CA125), and for postmenopausal, PI = -8.09 + 1.04 × ln(HE4) + 0.732 × ln(CA125). The cutoff points were defined at 11.4% for premenopausal and 29.9% for postmenopausal.

R-OPS was calculated using menopausal status, serum levels of CA125 (U/mL), HE4 (pM/L), and solid lesion: R-OPS = M × U × $(CA125 × HE4)^{1/2}$. The code of this index is 1 for premenopausal and 3 for postmenopausal women; U was coded as 1 for no solid lesion and 6 for presence of solid area. The cutoff point for R-OPS was set at 330.

Ultrasound scanning technique

All patients underwent preoperative transvaginal ultrasonography performed by the experienced gynecologists, gynecologic oncologists, or residency/fellowship trainees under supervision, adhering to the terms, definitions, and sonographic features standardized by the IOTA group. The transvaginal scans were done with a 4.0- to 12.0-MHz transducer, while the transabdominal scans used a 2.0- to 5.0-MHz transducer on the GE Voluson P8 system (GE Healthcare, Chicago, IL, USA).

Radioimmunoassay

Within 1 month before surgery, all patients were taken a blood sampling of 10 mL for CBC, blood urea nitrogen (BUN), creatinine (Cr), LFT, CA125, and HE4. Serum CA125 and HE4 were performed via an electrochemiluminescence immunoassay using Cobas e402/801 of Roche.
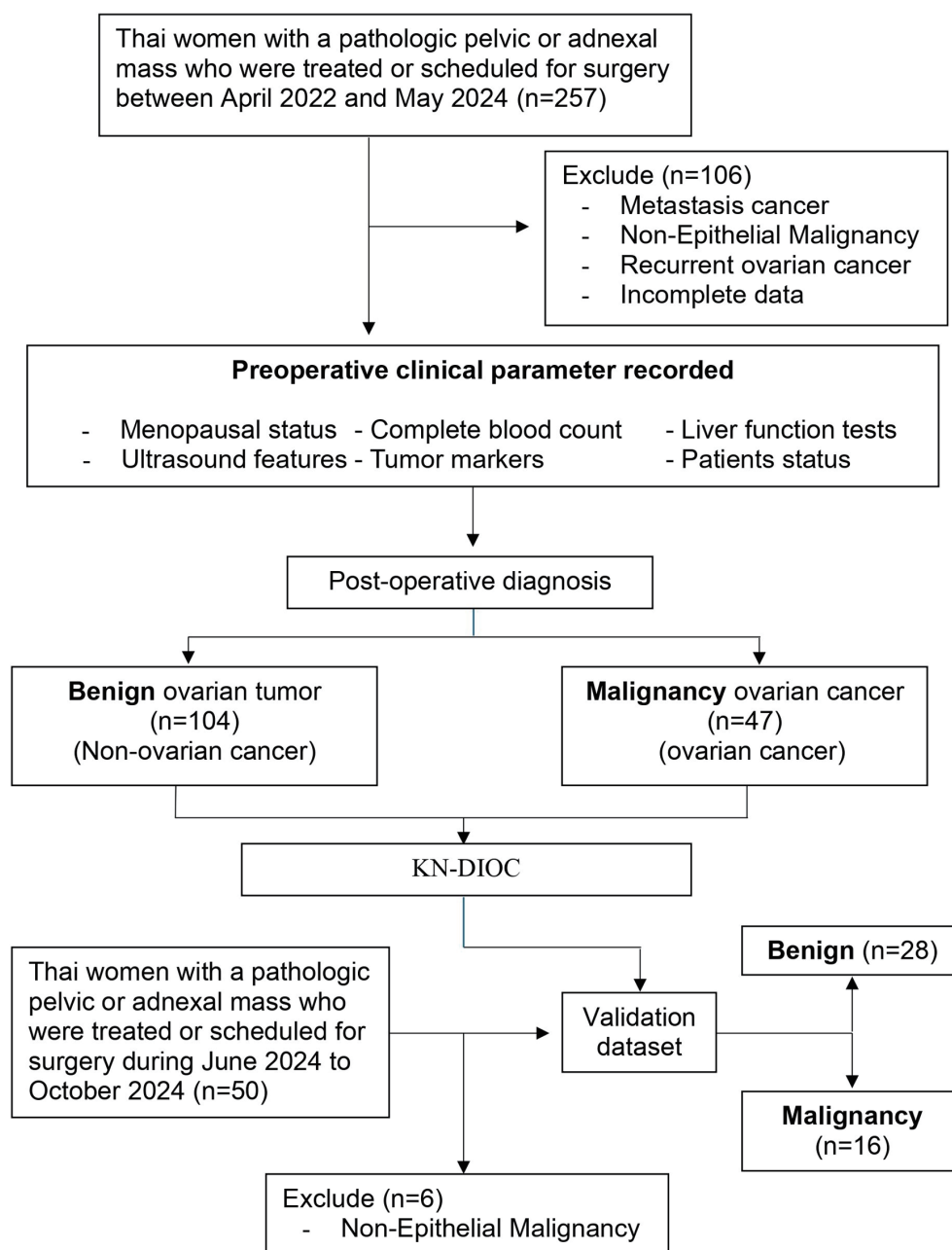
Histological or surgical diagnosis

The tumor specimens underwent histopathological examination at the Department of Anatomical Pathology in the same institute. The tumors were categorized into benign, borderline, or malignant based on the WHO classification system, and the malignant tumors were staged according to the International Federation of Gynecology and Obstetrics (FIGO) guidelines [11].

Statistical analyses and model development

All statistical analyses were performed using SPSS (IBM SPSS Statistics for Windows, version 26.0, released 2019, IBM Corp., Armonk, NY). Parametric continuous variables are presented as mean (standard deviation (SD)), while categorical variables are reported as percentages based on cytoreduction results. The clinical characteristics of participants were described separately for the development and validation datasets. Comparisons between benign and ovarian cancer patients within each dataset, as well as between the two datasets, were conducted using Student's t-test or Mann-Whitney U test, as appropriate, for continuous variables, and the Chi-square test for categorical variables.

In model development, we developed our preoperative diagnostic index using the development dataset by performing multivariate logistic regression analyses to identify significant predictors of ovarian cancer. Pearson's correlation was applied to evaluate relationships between predictors and detect multicollinearity problems within each group's predictors. Logistic regression analysis was conducted within each risk predictor group to identify statistically significant predictors associated with ovarian cancer, with each predictor being assessed using odds ratio (OR) and P-values. The preliminary analysis categorized the predictors into two groups: CBC and ultrasound features. Subsequently, significant predictors from each group were incorporated with CA125 levels and menopausal status to develop a preoperative diagnostic index using multivariate logistic regression. Generally, non-significant predictors were excluded from the analysis based on an OR of 1 and a P-value > 0.05. Regression analysis was then performed to obtain the coefficients for the remaining predictors. However, non-significant predictors may be retained in the analysis if deemed clinically relevant. The calibration of the logistic regression models was assessed using the Hosmer-Lemeshow goodness-of-fit statistic. The discriminative performance of the models was evaluated by calculating the area under the receiver operating characteristic curve (AUC-ROC) which is a graphical plot between sensitivity and 1 - specificity at various possible cutoff points. This study determined the optimal cutoff point by considering the point on the ROC curve that was closest to the perfect cutoff point (0, 1) [12].

Additionally, we report true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), from which sensitivity (Sn), specificity (Sp), accuracy (Acc), positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), negative likelihood ratio (LR-), and OR were calculated. Finally, the KN-DIOC

```
┌────────────────────────────────────────────────┐
│ Thai women with a pathologic pelvic or adnexal │
│ mass who were treated or scheduled for surgery │
│ between April 2022 and May 2024 (n=257)        │
└────────────────────────────────────────────────┘

                    ┌──────────────────────────────────┐
                    │ Exclude (n=106)                  │
                    │  -  Metastasis cancer            │
                    │  -  Non-Epithelial Malignancy    │
                    │  -  Recurrent ovarian cancer     │
                    │  -  Incomplete data              │
                    └──────────────────────────────────┘

┌────────────────────────────────────────────────────────────────────┐
│              Preoperative clinical parameter recorded              │
│                                                                    │
│  -  Menopausal status  - Complete blood count   - Liver function tests │
│  -  Ultrasound features - Tumor markers          - Patients status │
└────────────────────────────────────────────────────────────────────┘

                    ┌──────────────────────────┐
                    │ Post-operative diagnosis │
                    └──────────────────────────┘

┌────────────────────────────┐        ┌────────────────────────────┐
│   Benign ovarian tumor     │        │  Malignancy ovarian cancer │
│        (n=104)             │        │        (n=47)              │
│  (Non-ovarian cancer)      │        │    (ovarian cancer)        │
└────────────────────────────┘        └────────────────────────────┘

                    ┌──────────────────┐
                    │     KN-DIOC      │
                    └──────────────────┘

┌────────────────────────────┐                          ┌──────────────────────┐
│ Thai women with a pathologic│                          │  Benign (n=28)       │
│ pelvic or adnexal mass who  │     ┌──────────────┐     └──────────────────────┘
│ were treated or scheduled for│────▶│  Validation  │
│ surgery during June 2024 to │     │   dataset    │
│ October 2024 (n=50)         │     └──────────────┘     ┌──────────────────────┐
└────────────────────────────┘                          │  Malignancy          │
                                                         │   (n=16)             │
         ┌──────────────────────────────┐               └──────────────────────┘
         │ Exclude (n=6)                │
         │  -  Non-Epithelial Malignancy│
         └──────────────────────────────┘
```

Figure 1. Conceptual framework and patient flow chart for research and model development.

was validated using data from the validation dataset, and the AUC of ROC and its 95% confidence interval (CI) were recalculated.

Results

Clinical data from 307 patients were collected for potential inclusion in the development and validation of the preoperative diagnostic indices for EOC. We carefully screened and selected cases based on specific criteria, including data completeness, histopathological diagnosis, and clinical relevance. Of these, 257 patients were allocated to the development dataset and 50 to the validation dataset. Both datasets contained information on patient status, blood test results, biomarkers, ultrasound features, and other clinical data. Altogether, 106 patients and six patients were excluded from the development and validation datasets, respectively, due to non-epithelial malignancies, incomplete data, metastatic cancer, and recurrent ovarian cancer. The flow chart illustrating patient inclusion in the study is shown in Figure 1. A total of 195 patients were included in the final analysis. The development dataset included

Table 2. Logistic Regression Analysis of CBC and Ultrasound Features (Development Dataset)

Predictors	Coefficient	OR	95% CI	P-value*
CBC				
Platelets counts	0.009	1.009	1.005 - 1.013	< 0.001
Constant	-3.692	0.025	-	< 0.001
Ultrasound features				
Solid area	1.868	6.475	2.488 - 16.851	< 0.001
Papillary projection	2.395	10.974	4.117 - 29.250	< 0.001
Ascites	1.490	4.439	1.362 - 14.474	0.013
Constant	-2.757	0.063	-	< 0.001

*Significant at P < 0.05. CBC: complete blood count; CI: confident interval; OR: odds ratio.

151 patients, divided into 104 with benign ovarian tumor and 47 with ovarian cancer (prevalence: 31.13%). The validation dataset included 44 patients, divided into 28 with benign and 16 with ovarian cancer (prevalence: 36.36%). The clinical characteristics of both datasets are presented in Supplementary Material 2 (wjon.elmerpub.com). No significant differences were observed in clinical parameters between the two datasets. Similarly, there were no significant differences in ultrasound features or histopathologic diagnoses, specifically regarding tumor cysts, solid areas, and laterality. However, significant differences were found in the cyst wall, multilocular cyst, and ascites between the two datasets.

The clinical characteristics of the development dataset comparing benign ovarian tumors and ovarian cancer are presented in Supplementary Material 3 (wjon.elmerpub.com). Patients with ovarian cancer had significantly lower hemoglobin levels (11.34 vs. 11.99 g/dL, P = 0.012), and lymphocyte percentages (22.68% vs. 28.15%, P < 0.001). They also exhibited higher white blood cell counts (8,758.09 vs. 6,946.92 cells/mm^3, P = 0.010), neutrophil percentages (69.01% vs. 63.62%, P = 0.001), platelet counts (387.55 vs. 281.93 × 10^3 cells/mm^3, P < 0.001), neutrophil-to-lymphocyte ratio (NLR) (4.06 vs. 2.49, P < 0.001), and PLR (257.33 vs. 165.96, P < 0.001), but a lower LMR (4.18 vs. 5.50, P < 0.001). Serum CA125 levels were also significantly higher in patients with ovarian cancer (1,057.62 vs. 90.60 U/mL, P = 0.004). Regarding ultrasound features, ovarian cancer patients had significantly higher detection rates of solid areas (74.5% vs. 25.0%, P < 0.001), papillary projections (55.3% vs. 12.5%, P < 0.001), cyst wall irregularities (68.1% vs. 19.2%, P < 0.001), and ascites (31.9% vs. 6.7%, P < 0.001). Likewise, the clinical characteristics of the validation dataset are shown in Supplementary Material 4 (wjon.elmerpub.com). There were four significant parameters: basophil percentage (0.58% vs. 0.81%, P = 0.05), detected tumor size ≥ 7 cm (93.8% vs. 64.3%, P = 0.03), papillary projections (62.5% vs. 17.9%, P = 0.003), and ascites (62.5% vs. 10.7%, P < 0.001).

To evaluate the potential predictors for developing a preoperative diagnostic index, the analysis was divided into two groups based on CBC and ultrasound features. The potential predictors within each group were simultaneously analyzed using multivariate logistic regression analysis. One important consideration for logistic regression is the multicollinearity

problem, which occurs when independent variables in the regression model are highly correlated. To prevent multicollinearity, it is essential to assess the correlation between predictors using Pearson's or Spearman's correlation. Highly correlated predictors should then be either removed or combined to reduce multicollinearity before performing logistic regression analysis.

In the CBC group, hemoglobin showed a strong correlation with hematocrit (r = 0.958, P < 0.001). Neutrophils showed strong correlation with lymphocytes (r = -0.956, P < 0.001) and NLR (r = 0.830, P < 0.001). Similarly, lymphocytes showed strong correlation with NLR (r = -0.861, P < 0.001) and LMR (r = 0.773, P < 0.001). In this case, some predictors were eliminated from the analysis. The result of multivariate logistic regression analysis for CBC group is presented in Table 2. In contrast, individual predictors in ultrasound features group were not highly correlated, indicating no multicollinearity problem in this case. The results of multivariate logistic regression analysis are presented in Table 2. Finally, significant predictors included platelets counts, ultrasound finding of solid area, papillary projection, cyst wall, and ascites. These significant predictors, along with menopausal status, the presence of multilocular cyst, and CA125, will be used as initial predictors for developing the index with multivariate logistic regression analysis.

The final model, subsequently named the KN-DIOC, was developed by eliminating non-significant predictors from the multivariate logistic regression analysis. Consequently, platelet counts, serum CA125 levels, ultrasound findings of solid area, papillary projection, and multilocular cyst were included in the model. Additionally, menopausal status was retained. The results of the logistic regression analysis for the KN-DIOC are presented in Table 3, and the predictive index was described as follows:

$$\text{KN-DIOC} = P(Y=1) = P(EOC) = \frac{e^w}{1+e^w}$$

where W = -6.527 + 0.874 Menopausal + 0.008 Platelets + 1.484 Solid + 2.446 Papillary + 1.389 Multilocular + 0.002 CA125.

Based on the development dataset, the discriminative performance of the index was evaluated using the ROC curve, as shown in Figure 2, with an AUC of 0.926 (95% CI: 0.882

Table 3. Result of Logistic Regression Analysis for the KN-DIOC (Development Dataset)

Predictors	Coefficient	OR	95% CI	P-value*
Postmenopausal	0.874	2.396	0.795 - 7.219	0.121
Platelets counts	0.008	1.008	1.002 - 1.013	0.010
Solid area	1.484	4.410	1.449 - 13.423	0.009
Papillary projection	2.446	11.547	3.753 - 35.527	< 0.001
Multilocular cyst	1.389	4.009	1.264 - 12.715	0.018
CA125	0.002	1.002	1.001 - 1.004	0.007
Constant	-6.527	0.001	-	< 0.001

*Significant at P < 0.05. CA125: cancer antigen 125; CI: confident interval; KN-DIOC: KMUTT-NMU Diagnostic Index for Ovarian Cancer; OR: odds ratio.

- 0.970). Additionally, the P-value of the Hosmer-Lemeshow goodness-of-fit test was 0.293, indicating that the index model fit the data well. At the optimal cutoff point of 0.25, the index reported TP, FP, FN, and TN of 43, 14, 4, and 90, respectively.

The performance metrics (2 × 2 table) of the index were as follows: sensitivity was 91.5% (95% CI: 79.62 - 97.6), specificity was 86.5% (95% CI: 78.45 - 92.44), PPV was 75.4% (95% CI: 65.18 - 83.44), and NPV was 95.7% (95% CI: 89.78 - 98.29).
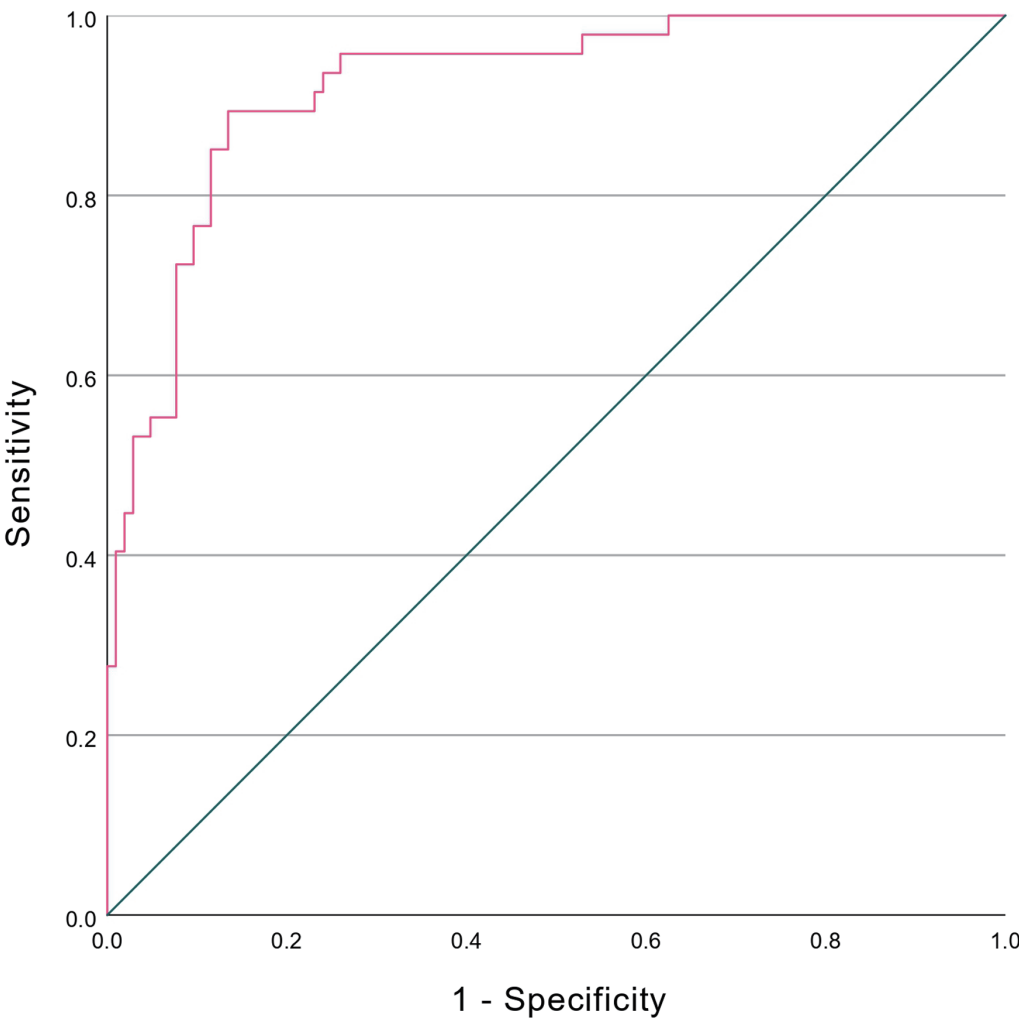
Figure 2. ROC curves of KN-IOC. KN-DIOC: KMUTT-NMU Diagnostic Index for Ovarian Cancer; ROC: receiver operating characteristic.

Table 4. Discriminative Performance of Preoperative Diagnosis Indices (Validation Dataset)

Test result indices	Cutoff point	AUC	P-value[a]	95% CI	P-value[b]	AUC difference
KN-DIOC	0.25	0.866	< 0.001	0.750 - 0.982	Ref.	Ref.
RMI I	200	0.853	< 0.001	0.720 - 0.986	0.857	0.013
RMI II	200	0.833	< 0.001	0.706 - 0.959	0.614	0.033
RMI III	200	0.853	< 0.001	0.731 - 0.974	0.846	0.013
RMI IV	450	0.857	< 0.001	0.744 - 0.971	0.881	0.009
ROMA	11.4/29.9[c]	0.900	< 0.001	0.795 - 1.000	0.558	-0.033
R-OPS	330	0.828	< 0.001	0.705 - 0.951	0.452	0.038

Significant at P < 0.05. [a]Null hypothesis: true area = 0.5. [b]Null hypothesis: true area difference = 0. [c]Premenopausal/postmenopausal cut-off points. AUC: area under the curve; CI: confident interval; KN-DIOC: KMUTT-NMU Diagnostic Index for Ovarian Cancer; OR: odds ratio; RMI: Risk of Malignancy Index; ROMA: Risk of Ovarian Malignancy Algorithm; R-OPS: Rajavithi-ovarian cancer predictive score.

The accuracy of the model was 88.1% (95% CI: 81.82 - 92.78). Additionally, the LR+ was 6.796 (95% CI: 4.14 - 11.15), the LR- was 0.098 (95% CI: 0.04 - 0.25), and the OR was 69.107 (95% CI: 21.47 - 222.47).

We evaluated the index's performance on the validation dataset using the optimal cutoff point. The results showed TP, FP, FN, and TN values of 15, 6, 1, and 22, respectively. The calculated performance metrics were sensitivity, specificity, accuracy, PPV, NPV, LR+, LR-, and OR of 93.75%, 78.57%, 84.09%, 71.43%, 95.65%, 4.375, 0.080, and 55, respectively. These values fell within the 95% CI obtained from the development dataset, indicating that the index's performance on the validation dataset is statistically consistent with that observed in the development dataset.

In addition, when comparing the performance of KN-DIOC with RMII-IV, ROMA, and R-OPS on the validation dataset, the discrimination performance and ROC curves are shown in Table 4 and Figure 3, with an AUC of 0.866 (95% CI: 0.750 - 0.982). All indices had significant ability to discriminate patients with and without ovarian cancer. The KN-DIOC demonstrated higher performance than RMII-IV and R-OPS. However, when comparing the discriminatory performance of the index with the other indices, there was no statistically significant difference in their ability to distinguish between patients with and without ovarian cancer. Nevertheless, when examining performance metrics based on the 2 × 2 table (Table 5), the index had the lowest FN (1 case) and the highest TP, resulting in the highest sensitivity, accuracy, NPV, LR-, and OR compared to the other indices.

Discussion

Ovarian cancer is a leading cause of gynecological cancer deaths worldwide, largely due to its late diagnosis and poor prognosis. The complexity of its presentation and the often-subtle symptoms that accompany early stages make it imperative to develop more precise diagnostic tools. Our study presents a novel preoperative diagnostic index that integrates CBC, tumor markers, and ultrasound features to improve diagnostic accuracy. The KN-DIOC exhibited a robust AUC of 0.866, indicating its excellent ability to differentiate between benign and malignant ovarian masses. The AUC is higher than those of RMI I-IV and R-OPS, and only slightly lower than ROMA, with no statistically significant differences in AUC, as shown in Table 4. However, a comparison of clinical performance indicators using 2 × 2 tables (Table 5) demonstrates the superior clinical performance of KN-DIOC. It achieved the highest sensitivity of 93.75%, with only one false-negative case among 16 patients with ovarian cancer. The highest NPV of 95.65% among all indices suggests a low probability of false-negative results, which is vital for ensuring that malignancies are not overlooked, indicating its strong capability to accurately identify patients without malignancy. In addition, its LR- was as low as 0.08, reinforcing the clinical confidence that a negative KN-DIOC result significantly reduces the probability of malignancy and may help avoid unnecessary operations. These properties suggest that KN-DIOC is particularly well suited to use as a "rule-out tool" in preoperative planning, especially in clinical contexts where missing a malignancy could have serious consequences. Although ROMA demonstrated a slightly higher AUC (0.9 vs. 0.866), KN-DIOC has advantages in terms of simplicity of use and does not require HE4, which remains limited in access in many settings, especially in resource-limited hospitals and countries. Therefore, KN-DIOC represents a practical and clinically valuable option for general medical practice, especially in real-world environments where diagnostic resources are constrained.

In the landscape of ovarian cancer diagnostics, traditional indices each present unique advantages and limitations. The RMI, extensively validated and recommended for clinical use, incorporates ultrasound features specific to ovarian cancer [2]. The RMI generally offers a sensitivity around 85% and a specificity near 97% when applied to Western populations [6]. However, the performance of RMI tends to diminish in Asian cohorts, where the sensitivity frequently falls due to variations in tumor biology that affect biomarker expressions differently than in Western populations [7]. Although the KN-DIOC model was developed specifically for Thai patients, its reliance on routine clinical parameters makes it potentially applicable to other Southeast Asian populations with similar resource constraints and ovarian cancer profiles. Several Southeast Asian countries share comparable demographic, clinical, and histo-
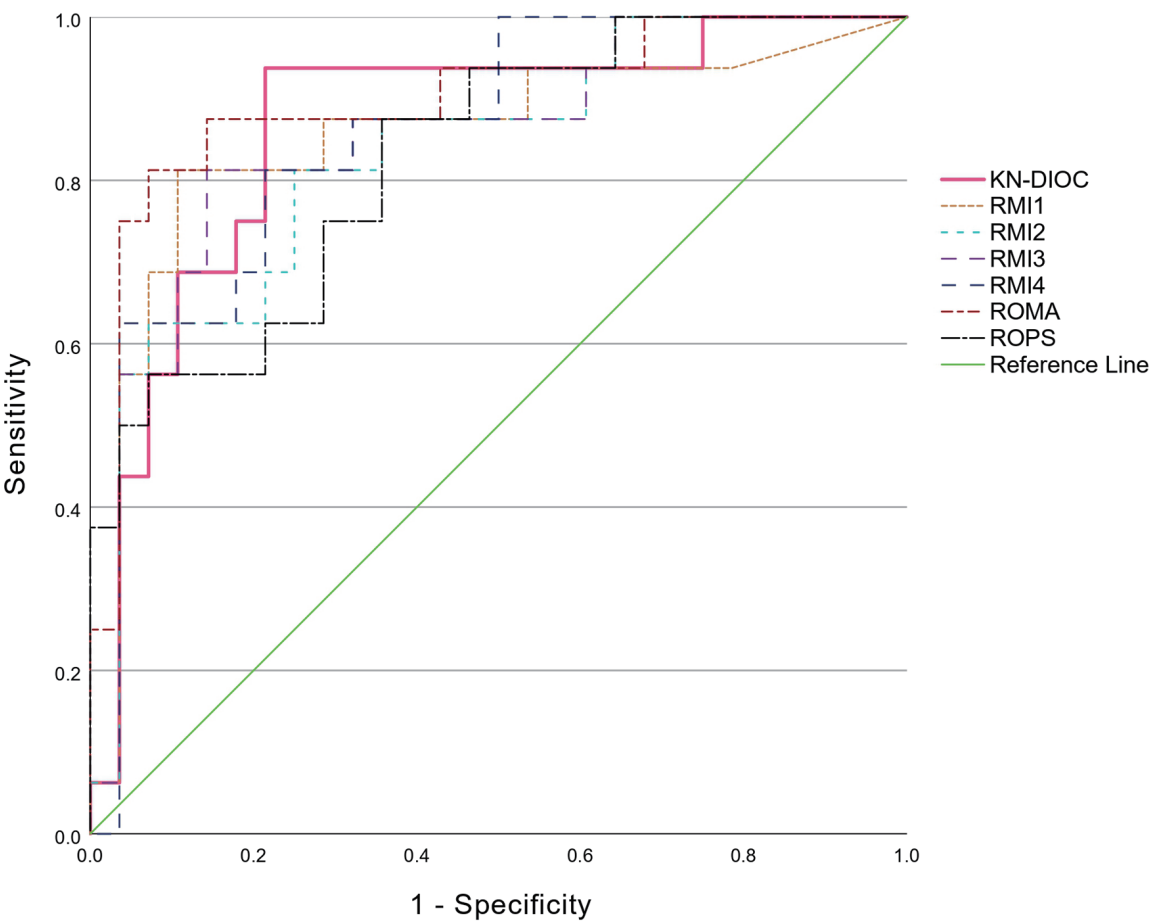
Figure 3. Comparison of ROC curves: KN-IOC, RMI I-IV, ROMA, and R-OPS. KN-DIOC: KMUTT-NMU Diagnostic Index for Ovarian Cancer; RMI: Risk of Malignancy Index; ROC: receiver operating characteristic; ROMA: Risk of Ovarian Malignancy Algorithm; R-OPS: Rajavithi-Ovarian Cancer Predictive Score.

Table 5. Diagnostic Test Performance Metrics Based on Validation Dataset

	KN-DIOC	RMI I	RMI II	RMI III	RMI IV	ROMA	R-OPS
TP	15	11	13	11	11	13	10
FP	6	2	9	3	6	4	7
FN	1	5	3	5	5	3	6
TN	22	26	19	25	22	24	21
Sn (%)	93.75	68.75	81.25	68.75	68.75	81.25	62.5
Sp (%)	78.57	92.86	67.86	89.29	78.57	85.71	75.0
Acc (%)	84.09	84.09	72.73	81.82	75.0	84.09	70.45
PPV (%)	71.43	84.62	59.09	78.57	64.71	76.47	58.82
NPV (%)	95.65	83.87	86.36	83.33	81.48	88.89	77.78
LR+	4.375	9.625	2.528	6.417	3.208	5.688	2.5
LR-	0.080	0.337	0.276	0.35	0.398	0.219	0.5
OR	55	28.6	9.15	18.33	8.067	26	5

Acc: accuracy; FN: false negatives; FP: false positives; KN-DIOC: KMUTT-NMU Diagnostic Index for Ovarian Cancer; LR+: positive likelihood ratio; LR-: negative likelihood ratio; NPV: negative predictive value; OR: odds ratio; PPV: positive predictive value; RMI: Risk of Malignancy Index; ROMA: Risk of Ovarian Malignancy Algorithm; R-OPS: Rajavithi-ovarian cancer predictive score; Sn: sensitivity; Sp: specificity; TN: true negatives; TP: true positives.

pathological characteristics of ovarian cancer. For instance, clear cell and endometrioid subtypes are notably more prevalent in East and Southeast Asia compared to Western populations [13]. Further validation studies are thus encouraged to explore the applicability and potential recalibration of the KN-DIOC model in other regional populations.

Similarly, the ROMA typically demonstrates sensitivities between 75% and 85% and specificities from 90% to 95% [14]. Although ROMA is quite effective, its reliance solely on specific biomarkers such as CA125 and HE4 without accounting for systemic inflammatory responses can restrict its diagnostic reach. This limitation becomes particularly notable in patients with atypical symptoms or those in early stages of tumor development, where traditional biomarkers might not yet be elevated.

Contrastingly, the ADNEX model, which relies heavily on detailed ultrasound data and requires substantial expertise for the accurate interpretation of complex morphological features, might not be as readily applicable in settings with limited resources or less specialized ultrasound expertise [15].

Ultrasound, while indispensable in ovarian cancer diagnosis, often requires high expertise and quality equipment to discern subtle features associated with malignancy. Reducing the reliance on complex ultrasound criteria and enhancing it with CBC data leverages the strengths of both modalities. For instance, elevated platelet counts (thrombocytosis) have been consistently associated with worse outcomes in cancer patients and are emerging as a significant biomarker in ovarian cancer [16]. The KN-DIOC, therefore, incorporates platelet counts alongside traditional markers like CA125 and HE4, providing a multi-faceted approach to diagnosis.

The major strength of the index lies in its comprehensive approach, combining various diagnostic modalities to offer a robust tool that adapts to the specific needs of the Thai population. However, the study's limitations include its single-center design and a relatively small validation dataset, comprising only 16 patients with ovarian cancer, which may limit the strength of external validation and affect the generalizability of the findings. Future research should aim to validate these findings across multiple centers and diverse populations within Thailand and other Asian countries. Expanding the scope of the index to include genetic markers and advanced imaging techniques could also provide deeper insights and refine its accuracy further.

Conclusion

This study has developed a novel preoperative diagnostic index tailored specifically for Thai patients, integrating menopausal status, specific ultrasound findings, platelet counts, and serum CA125 levels. This unique combination harnesses the most current and relevant patient data, enhancing its applicability and effectiveness within Thai and potentially other Asian populations. The KN-DIOC has demonstrated robust performance metrics, indicating strong potential for practical application in clinical settings to aid general physicians and gynecologists in assessing ovarian tumors. It promises to support more accurate management planning and patient prioritization for surgery, thereby improving outcomes in ovarian cancer care.

Given its strong performance and the integration of easily accessible clinical data, this index represents a significant advancement over existing diagnostic tools by offering a more adaptable and practical approach for settings with varying levels of resources and expertise.

Supplementary Material

Suppl 1. Components and Score Rating of Risk of Malignancy Index I-IV.
Suppl 2. Clinical Characteristics Between Datasets.
Suppl 3. Clinical Characteristics in Development Dataset.
Suppl 4. Clinical Characteristics in Validation Dataset.

Acknowledgments

Financial Disclosure

Conflict of Interest

None to declare.

Informed Consent

Not applicable.

Author Contributions

Conceptualization: ST; methodology: ST and TS; software: ST; validation: ST, TS, and WC; formal analysis: ST; investigation: ST and TS; resources: ST and WC; data curation: ST and WC; writing - original draft preparation: ST; writing - review and editing: ST, TS, and WC; visualization: ST; supervision: TS. All authors have read and agreed to the published version of the manuscript.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Abbreviations

Acc: accuracy; ACOG: American College of Obstetricians and Gynecologists; AUC: area under the curve; CA125: cancer antigen 125; CA19-9: cancer antigen 19-9; CBC: complete blood count; CCC: clear cell carcinoma; CI: confidence interval; EOC: epithelial ovarian cancer; FIGO: International Federation of Gynecology and Obstetrics; FN: false negatives; FP: false positives; GCTs: germ cell tumors; HDI: Human Development Index; HE4: human epididymal protein 4; HGSC: high-grade serous carcinoma; IOTA: International Ovarian Tumor Analysis; KN-DIOC: KMUTT-NMU Diagnostic Index for Ovarian Cancer; LMR: lymphocyte-to-monocyte ratio; LR+: positive likelihood ratio; LR-: negative likelihood ratio; M: menopausal status; NPV: negative predictive value; OR: odds ratio; PLR: platelet-to-lymphocyte ratio; PPV: positive predictive value; RMI: Risk of Malignancy Index; ROC: receiver operating characteristic; ROMA: Risk of Ovarian Malignancy Algorithm; R-OPS: Rajavithi-ovarian cancer predictive score; S: tumor size score; SD: standard deviation; Sn: sensitivity; Sp: specificity; TN: true negatives; TP: true positives; U: ultrasound score

References

1. International Agency for Research on Cancer. Latest global cancer data. GLOBOCAN database 2020. Lyon: IARC; 2020 [cited Feb 13, 2025]. Available from: https://www.iarc.who.int/news-events/latest-global-cancer-data-globocan-database-2020/.
2. 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(10):922-929. doi pubmed
3. Moore RG, McMeekin DS, Brown AK, DiSilvestro P, Miller MC, Allard WJ, Gajewski W, et al. A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. Gynecol Oncol. 2009;112(1):40-46. doi pubmed
4. Timmerman D, Testa AC, Bourne T, Ameye L, Jurkovic D, Van Holsbeke C, Paladini D, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol. 2008;31(6):681-690. doi pubmed
5. Tingulstad S, Hagen B, Skjeldestad FE, Onsrud M, Kiserud T, Halvorsen T, Nustad K. Evaluation of a risk of malignancy index based on serum CA125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses. Br J Obstet Gynaecol. 1996;103(8):826-831. doi pubmed
6. Tingulstad S, Hagen B, Skjeldestad FE, Halvorsen T, Nustad K, Onsrud M. The risk-of-malignancy index to evaluate potential ovarian cancers in local hospitals. Obstet Gynecol. 1999;93(3):448-452. pubmed
7. Yamamoto Y, Yamada R, Oguri H, Maeda N, Fukaya T. Comparison of four malignancy risk indices in the preoperative evaluation of patients with pelvic masses. Eur J Obstet Gynecol Reprod Biol. 2009;144(2):163-167. doi pubmed
8. Chirdchim W, Wanichsetakul P, Phinyo P, Patumanond J, Suwannarurk K, Srisomboon J. Development and validation of a predictive score for preoperative diagnosis of early stage epithelial ovarian cancer. Asian Pac J Cancer Prev. 2019;20(4):1207-1213. doi pubmed
9. Moolthiya W, Yuenyao P. The risk of malignancy index (RMI) in diagnosis of ovarian malignancy. Asian Pac J Cancer Prev. 2009;10(5):865-868. pubmed
10. Yanaranop M, Tiyayon J, Siricharoenthai S, Nakrangsee S, Thinkhamrop B. Rajavithi-ovarian cancer predictive score (R-OPS): A new scoring system for predicting ovarian malignancy in women presenting with a pelvic mass. Gynecol Oncol. 2016;141(3):479-484. doi pubmed
11. Prat J, Oncology FCoG. Staging classification for cancer of the ovary, fallopian tube, and peritoneum. Int J Gynaecol Obstet. 2014;124(1):1-5. doi pubmed
12. Pepe MS. The statistical evaluation of medical tests for classification and prediction, Oxford University Press, Oxford. 2003.
13. Momenimovahed Z, Tiznobaik A, Taheri S, Salehiniya H. Ovarian cancer in the world: epidemiology and risk factors. Int J Womens Health. 2019;11:287-299. doi pubmed
14. Moore RG, Miller MC, Disilvestro P, Landrum LM, Gajewski W, Ball JJ, Skates SJ. Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. Obstet Gynecol. 2011;118(2 Pt 1):280-288. doi pubmed
15. Van Calster B, Timmerman D, Bourne T, Testa AC, Van Holsbeke C, Domali E, Jurkovic D, et al. Discrimination between benign and malignant adnexal masses by specialist ultrasound examination versus serum CA-125. J Natl Cancer Inst. 2007;99(22):1706-1714. doi pubmed
16. Wang L, Wang X, Guo E, Mao X, Miao S. Emerging roles of platelets in cancer biology and their potential as therapeutic targets. Front Oncol. 2022;12:939089. doi pubmed