



Distinguishing benign from malignant pelvic mass utilizing an algorithm with HE4, menopausal status, and ultrasound findings

Sarikapan Wilailak¹, Karen KL Chan², Chi-An Chen³, Joo-Hyun Nam⁴, Kazunori Ochiai⁵, Tar-Choon Aw⁶, Subathra Sabaratnam⁷, Sudarshan Hebbar^{8,*}, Jaganathan Sicksan⁸, Beth A Schodin⁸, Chuenkamon Charakorn¹, Walfrido W Sumpaico⁹

¹Department of Obstetrics and Gynecology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; ²Department of Obstetrics and Gynecology, University of Hong Kong, Hong Kong; ³Department of Obstetrics and Gynecology, National Taiwan University Hospital, Taipei, Taiwan; ⁴Department of Obstetrics and Gynecology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea; ⁵Department of Obstetrics and Gynecology, The Jikei University, Tokyo, Japan; ⁶Department of Laboratory Medicine, Changi General Hospital, Singapore; ⁷Department of Pathology, Hospital Sultanah Aminah, Johor Bahru, Johor, Malaysia; ⁸Abbott Diagnostics, Abbott Park, IL, USA; ⁹Department of Obstetrics and Gynecology, MCU-FDT Medical Foundation, Caloocan, Philippines

Objective: The purpose of this study was to develop a risk prediction score for distinguishing benign ovarian mass from malignant tumors using CA-125, human epididymis protein 4 (HE4), ultrasound findings, and menopausal status. The risk prediction score was compared to the risk of malignancy index and risk of ovarian malignancy algorithm (ROMA).

Methods: This was a prospective, multicenter (n=6) study with patients from six Asian countries. Patients had a pelvic mass upon imaging and were scheduled to undergo surgery. Serum CA-125 and HE4 were measured on preoperative samples, and ultrasound findings were recorded. Regression analysis was performed and a risk prediction model was developed based on the significant factors. A bootstrap technique was applied to assess the validity of the HE4 model.

Results: A total of 414 women with a pelvic mass were enrolled in the study, of which 328 had documented ultrasound findings. The risk prediction model that contained HE4, menopausal status, and ultrasound findings exhibited the best performance compared to models with CA-125 alone, or a combination of CA-125 and HE4. This model classified 77.2% of women with ovarian cancer as medium or high risk, and 86% of women with benign disease as very-low, low, or medium-low risk. This model exhibited better sensitivity than ROMA, but ROMA exhibited better specificity. Both models performed better than CA-125 alone.

Conclusion: Combining ultrasound with HE4 can improve the sensitivity for detecting ovarian cancer compared to other algorithms.

Keywords: Algorithms; CA-125 Antigen; Ovarian Neoplasms; Prospective Studies; Regression Analysis; Sensitivity and Specificity

Received July 1, 2014, Revised Sep 24, 2014, Accepted Oct 6, 2014

*Current affiliation: Reata Pharmaceuticals, Irving, TX, USA

Correspondence to Sarikapan Wilailak

Department of Obstetrics and Gynecology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand. E-mail: sarikapan.wil@mahidol.ac.th

Copyright © 2015. Asian Society of Gynecologic Oncology, Korean Society of Gynecologic Oncology

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

In 2009, 215,500 new cases of ovarian cancer were diagnosed worldwide and 146,000 women died of the disease [1]. Among gynecologic malignancies, ovarian cancer remains the most lethal. Unfortunately, ovarian cancer does not have a precancerous lesion like cervical cancer. Therefore, the aim for better outcome is limited to an attempt to detect early-stage disease, when the symptoms are often nonspecific but the prognosis is good, or to distinguish between benign and malignant pelvic masses.

Studies have shown that patients with ovarian cancer have improved outcomes when they are managed in specialized centers by gynecologic oncologists [2]. Gynecologic oncologists are trained to perform the difficult cytoreduction required for optimal outcome, but less than half of the women that are ultimately diagnosed with ovarian cancer are referred to a gynecologic oncologist or specialist [3]. To increase the number of women with ovarian cancer that are referred to a specialist, risk prediction scores have been developed to identify women that have a high risk of ovarian cancer by incorporating clinical data (e.g., menopausal status), imaging (e.g., ultrasound finding) and/or tumor markers (e.g., cancer antigen 125 [CA-125], human epididymis protein 4 [HE4]) in regression models. Examples of risk prediction models include: (1) risk of malignancy index (RMI) which uses ultrasound findings, menopausal status, and CA-125 to identify women at high risk of having a malignant ovarian tumor [4]; (2) risk of ovarian malignancy algorithm (ROMA), which uses menopausal status plus CA-125 and HE4 values [5]; (3) OVA1, which uses CA-125, 2-microglobulin, apolipoprotein A1, transthyretin, and transferrin [6]; and (4) LR2, an ultrasound prediction model developed by the International Ovarian Tumor Analysis (IOTA) study [7]. Except for the ultrasound prediction model, all of the other models include the tumor marker CA-125, which is elevated in over 80% of advanced ovarian cancers. However, CA-125 is also elevated in non-malignant conditions, such as endometriosis and fibroids, which can decrease the specificity and positive predictive value (PPV) for ovarian cancer [8].

HE4 has been identified as a marker for ovarian cancer with improved specificity over CA-125 [9]. It is logical to determine if an algorithm can be developed using HE4, clinical data, and ultrasound features. HE4 is a protein present in high concentrations in the male epididymis and has also been found in the serum of patients with ovarian cancer [10]. The function of the protein is unknown, but it has been shown to be sensitive and specific for ovarian cancer [11]. The ROMA was therefore developed by considering both HE4 and CA-125 tumor marker concentrations in the serum plus menopausal status [12].

ROMA has been evaluated in several studies and has shown good performance in pre- and postmenopausal women for distinguishing between benign and malignant pelvic mass [11-15]. ROMA has also been shown to exhibit similar or better discrimination of cancer from benign tumors than the RMI [16,17]. The ultrasound prediction model developed by IOTA has been shown to have better diagnostic performance than ROMA in the hands of expert sonographers [18]. The objective of this study is to determine if a combination of CA-125/HE4, ultrasound features and menopausal status in a logistic regression model could improve the prediction of ovarian cancer in women with a pelvic mass compared to ROMA or RMI.

The study population was previously described in "The use of HE4 in the prediction of ovarian cancer in Asian women with a pelvic mass," published by Chan et al. [19] in 2013. The original study concluded that ROMA showed similar sensitivity to CA-125, but improved specificity and PPV. Our current study will determine if the sensitivity can be improved by adding the ultrasound information.

MATERIALS AND METHODS

We used an existing dataset from a prospective multicenter cohort involving six centers in Asian countries. Of the 414 women enrolled in the original study, 328 had documented ultrasound information that was used in this study. Details of the study design and methods are clearly described in Chan et al. [19]. Briefly, consecutive women (age 18 years or older) diagnosed with an adnexal mass were enrolled. Cases were ineligible if they had met following criteria: history of ovarian or primary peritoneal cancer, any known malignancy, or history of bilateral oophorectomy. The study was approved by the institutional review boards at each site. Written informed consent was obtained from all subjects. Blood samples were collected prior to surgery in serum separator tubes and were centrifuged, aliquoted and frozen within 4 hours. The samples were stored at -20°C or colder at the individual study sites and were shipped on dry ice at the end of sample accrual to the central testing laboratory at Changi General Hospital, Singapore. Two representative H&E stained slides of the final adnexal pathology for each subject were sent to the central pathology laboratory at Hospital Sultanah Aminah, Johor Bahru, Johor, Malaysia for review. At the central laboratory, all samples were tested using the ARCHITECT CA-125 II, HE4, carcinoembryonic antigen (CEA), and follicle stimulating hormone (FSH) assays (Abbott Diagnostics, Abbott Park, IL, USA) according to the manufacturer's instructions. If the self-

reported menopausal status was not available from the case report form, the women's age and FSH values were used to assign the menopausal status. In addition, demographic and clinical data including age, menopausal status, previous CA-125 values (if known), histopathological diagnosis, surgical diagnosis, and stage of disease were collected and entered into a case report form. Ultrasound findings were also documented as the following 5 features: multiloculated, solid nodule, bilaterality, ascites, peritoneal metastases.

The RMI and ROMA were calculated following the suggestion of Jacobs and Moore, respectively [4,5]. Data were described using mean and frequency for continuous and categorical data, respectively. A bivariate logistic regression was applied to assess association between risk factor and cancer. Variables with $p < 0.15$ from this step were simultaneously considered in the multivariate logistic model. A likelihood ratio test was applied to select variable in the model. Goodness of fit of the final model was then assessed using Hosmer-Lemeshow chi-square. A receiver operating characteristic (ROC) curve analysis was applied to assess performance of the final model. The C statistic along with its 95% confidence interval (CI) for our model and the original ROMA and RMI models were calculated and compared. The net reclassification improvement (NRI) was applied to assess improvement in classifying of the HE4 model when compared with the ROMA [20].

Internal validation of the risk prediction score of the HE4 model was assessed using a bootstrap technique with 200 replications [21,22]. For each replication, the HE4 model was constructed including HE4, menopausal status, and ultrasound findings; risk prediction scores were calculated based on

estimated coefficients, and finally predictive performance parameters (i.e., predicted probability and the C statistic) were estimated. The Somer'D correlation was applied to assess association between the observed and predicted values of ovarian cancer, called D_{boot} . Calibration coefficient was then estimated by subtracting the original Somer'D correlation coefficient

Table 1. Characteristic of patients (n=328)

Characteristic	Value
Age (yr)	41.2±13.0
Menopausal status	
Premenopause	251 (76.5)
Postmenopause	77 (23.5)
Number of US features	
None	125 (38.1)
1	128 (39.0)
2	55 (16.8)
3	13 (4.0)
4	7 (2.1)
CA-125 (U/mL)	23.85 (2.5–1,000)
HE4 (pmol/L)	35 (16.7–1,500)
FSH (mIU/mL)	5.4 (0.1–119.0)
CEA (ng/mL)	1.4 (0.5–216.4)

Values are presented as mean±SD, number (%), or median (range). US features include multiloculated, solid nodule, bilaterality, ascites, peritoneal metastases. CA-125, cancer antigen 125; CEA, carcinoembryonic antigen; FSH, follicle-stimulating hormone; HE4, human epididymis protein 4; US, ultrasound.

Table 2. Comparison of patients characteristics between cancer and benign groups

Characteristic	Group		p-value	OR (95% CI)
	Cancer (n=57)	Benign (n=271)		
CA-125 (U/mL), median (range)*	130.4 (4.8–1,000)	20 (2.5–1,000)	<0.001	1.01 (1.00–1.02)
HE4 (pmol/L), median (range)*	99.2 (24.9–1,500)	33.4 (16.7–152.9)	<0.001	1.06 (1.04–1.07)
Age (yr)	51.12±13.37	39.13±11.97	<0.001	1.07 (1.05–1.10)
Menopausal status				
Premenopause	24±42.86	227±83.76		1
Postmenopause	33±57.14	44±16.24	<0.001	1.96 (1.34–2.58)
Number of US features				
None	6±10.53	119±43.91		1
1	16±28.07	112±41.33	0.036	2.83 (1.07–7.50)
2	16±28.07	39±14.39	<0.001	8.14 (2.98–22.24)
≥3	19±33.33	1±0.37	<0.001	376.83 (42.95–3,305.94)

Values are presented as mean±SD compare means using t-test. US features include multiloculated, solid nodule, bilaterality, ascites, peritoneal metastases. CA-125, cancer antigen 125; HE4, human epididymis protein 4; US, ultrasound. *Compared medians using Mann-Whitney rank test.

with the mean D_{boot} . Discriminative performance was assessed by comparing the original C statistic with an average C statistic from the bootstraps. All analyses were performed using STATA ver.12.0 (StataCorp, College Station, TX, USA). A p-value less than 0.05 was considered as statistically significant.

RESULTS

Four-hundred and fourteen patients were enrolled from

Korea (n=170), Thailand (n=121), the Philippines (n=57), Hong Kong (n=24), Taiwan (n=24), and Japan (n=18) as described in the original study [19]. Complete ultrasound data were available for 328 patients (79.2%) enrolled in that study and these were included in the current analyses. As described in **Table 1**, mean age was 41.2 ± 13.0 years; 251 (76.5%) were premenopausal and 77 women (23.5%) were postmenopausal. About 62% of patients had one or more features from ultrasound findings. Median CA-125 and HE4 were 23.9 U/mL (range, 2.5 to 1,000 U/mL) and 35 pmol/L (range, 16.7 to 1,500), respec-

Table 3. Association between HE4, CA-125, other risk factors in ovarian cancer

Risk factor	Coefficient	SE	t	p-value	95% Confidence interval
HE4 (pmol/L)	0.04	0.01	4.013	<0.001	0.02 to 0.06
Menopausal status					
Premenopause	0.00				
Postmenopause	0.82	0.49	1.658	0.097	-0.15 to 1.79
Number of US features					
None	0.00				
1	0.50	0.57	0.883	0.377	-0.61 to 1.61
2	1.68	0.57	2.961	0.003	0.57 to 2.80
≥3	3.47	1.26	2.750	0.006	1.00 to 5.94
Constant	-4.87	0.61	-7.99	<0.001	-6.06 to -3.67
CA-125 (U/mL)	0.003	0.001	3.350	0.001	0.001 to 0.005
Menopause					
Premenopause	0.00	0.00			0
Postmenopause	1.63	0.39	4.140	<0.001	0.86 to 2.40
Number of US features					
None	0.00	0.00			0
1	1.02	0.52	1.962	0.050	0.00 to 2.03
2	1.67	0.56	3.007	0.003	0.58 to 2.77
≥3	4.45	1.16	3.837	0.000	2.18 to 6.72
Constant	-3.70	0.48	-7.74	<0.001	-4.64 to -2.76
HE4 & CA-125					
HE4 (pmol/L)	0.04	0.01	3.892	<0.001	0.02 to 0.06
CA-125 (U/mL)	0.0003	0.00	0.295	0.768	-0.001 to -0.002
Menopause					
Premenopause	0.00	0.00			0
Postmenopause	0.82	0.49	1.661	0.097	-0.15 to -1.79
Number of US features					
None	0.00	0.00			0
1	0.51	0.57	0.894	0.371	-0.60 to -1.62
2	1.66	0.58	2.865	0.004	0.52 to 2.79
≥3	3.37	1.29	2.605	0.009	0.83 to 5.90
Constant	0.008	0.005	-7.990	<0.001	0.002 to 0.023

US features include multiloculated, solid nodule, bilaterality, ascites, and peritoneal metastases.

CA-125, cancer antigen 125; HE4, human epididymis protein 4; SE, standard error of the mean; US, ultrasound.

tively. The incidence of ovarian cancer was 17.3% (95% CI, 13.3 to 21.5) with the majority being epithelial ovarian cancers in both the pre- and postmenopausal groups (n=22 and 43, respectively). Out of the 65 epithelial ovarian cancers, there were 17 women (26.2%) with stage I disease, 5 (7.7%) with stage II, 26 (40.0%) with stage III and 11 (16.9%) with stage IV disease.

A bivariate logistic regression was applied and suggested that HE4, CA-125, age, menopausal status, and ultrasound finding were significantly associated with ovarian cancer (Table 2). Three multiple logistic models were constructed by including each studied marker (i.e., HE4, CA-125, and

HE4+CA-125) and covariables (age, menopausal status, and ultrasound finding) in the model. A LR test was applied and suggested that all variables except age were significantly associated with ovarian cancer (Table 3). These models fitted well with the data as per the Hosmer-Lemeshow goodness of fit of 12.6 (p=0.169), 4.9 (p=0.765), and 11.4 (p=0.181) for HE4, CA-125, and HE4+CA-125 models, respectively. Coefficients of each model were then used to calculate scores for individual patients. The ROC analysis was applied and suggested that the C-statistics for these corresponding models were respectively 0.893 (95% CI, 0.837 to 0.949), 0.865 (95% CI, 0.804 to 0.926), and 0.893 (95% CI, 0.837 to 0.949) (Fig. 1). This indicated that the HE4 model was significantly better in discriminative ability than the CA-125 model (p=0.009); whereas adding CA-125 in the model that contained HE4 did not significantly improve discriminative ability (p=0.897) when compared with the HE4 model. In addition, the HE4 model was further compared with the RMI and ROMA models. This showed that the C statistic for HE4 model was significantly different when compared with the RMI (p=0.020) but not for the ROMA (p=0.118).

We further compared performance in classification between our HE4 and ROMA models using NRI statistics (Table 4). Based on the reference model of ROMA, the estimated probability of having ovarian cancer was divided into 4 groups according to quartile distribution with the cutoffs of <0.035, 0.035–0.060, 0.060–0.123, and >0.123, respectively. Applying these cutoffs to classify probability estimated by the HE4 model led us to assess reclassification improvements in cancer and benign groups. As described in Table 4, the green shade referred to perfect agreement between the 2 scores, the blue and orange

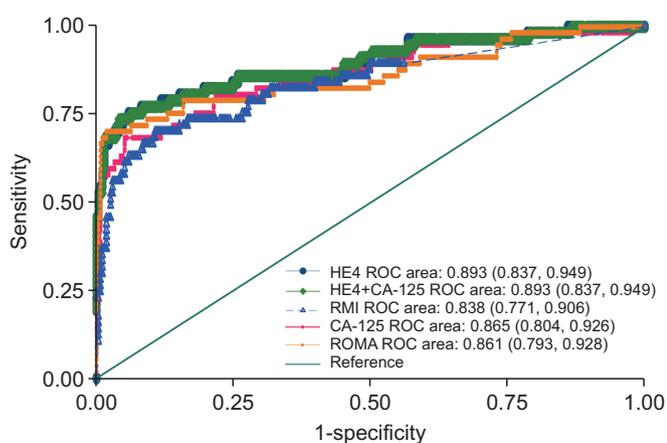


Fig. 1. Receive operation characteristic curve plots for risk prediction models: human epididymis protein 4 (HE4), cancer antigen 125 (CA-125), HE4+CA-125, risk of malignancy index (RMI), risk of ovarian malignancy algorithm (ROMA).

Table 4. Net reclassification improvement of the HE4 model compared with the ROMA

Cancer group	HE4 model				Total
	<0.035	0.035–0.060	0.060–0.123	>0.123	
ROMA					
<0.035	1	2	1	1	5
0.035–0.060	1	3	0	1	5
0.060–0.123	0	0	2	2	4
>0.123	0	1	0	42	43
Total	2	6	3	46	57
ROMA benign group					
<0.035	51	16	9	1	77
0.035–0.060	29	34	6	8	77
0.060–0.123	10	37	17	14	78
>0.123	0	4	17	18	39
Total	90	91	49	41	271

HE4, human epididymis protein 4; ROMA, risk of ovarian malignancy algorithm.

Table 5. Simplified score and its performance

Score cut-off	Risk of cancer	Group		Sensitivity (%)	Specificity (%)	PV+ (%)	LR+ (95% CI)
		Cancer	Benign				
<1.49	Very low	2	79				
≥1.49	Low	6	77	96.5	29.2	22.3	1.36 (1.24–1.49)
≥1.94	Low-medium	4	77	86.0	57.6	29.9	2.03 (1.70–2.41)
≥2.95	Medium	1	16	78.9	86	54.2	5.63 (4.07–7.78)
≥3.33	High	44	22	77.2	91.9	66.7	9.51 (6.22–14.50)

CI, confidence interval; LR+, likelihood ratio positive; PV+, positive predictive value.

shades referred to improved classification by our model and ROMA, respectively. The reclassification improvements were 8.8% (i.e., $[(2+1+1+1+2)-(1+1)]/57$) in the cancer group and 15.9% (i.e., $[(29+10+37+4+17)-(16+9+1+6+8+14)]/271$) in the benign groups. This indicated that our HE4 model could improve classification of cancer and non-cancer by 8.8% and 15.9% when compared with ROMA. The estimated overall NRI was 24.7% (95% CI, -39.8 to 89.1), but this was not significant from 0.

For applying the score in clinical practice, a scoring scheme from the HE4 model can be calculated using coefficient described in **Table 3** as follows:

$$\begin{aligned} \text{Score} = & 0.04 \times \text{HE4} + 0.82 \times (\text{MS} = \text{postmenopause}) + \\ & [0 \text{ (US feature number} = 0) \text{ or} \\ & 0.5 \text{ (US feature number} = 1) \text{ or} \\ & 1.68 \text{ (US feature number} = 2) \text{ or} \\ & 3.47 \text{ (US feature number} \geq 3)] \end{aligned}$$

Every actual value of HE4 level was multiplied by 0.04. Menopausal status (MS) was coded as 1 and scored as 0.82 for postmenopausal women and 0 for premenopausal women; ultrasound features were coded as 0, 1, 2, and ≥ 3 and dummy variables were created using 0 feature as the reference, then scored as 0, 0.5, 1.68, and 3.47, respectively. For ease of use and simplicity, the estimated score was classified into 5 groups according to score's distribution and its performance, scores of <1.49 , ≥ 1.49 , ≥ 1.94 , ≥ 2.95 , and ≥ 3.33 corresponded to very low, low, low-medium, medium, and high risk of having ovarian cancer with positive likelihood ratios of 1, 1.36, 2.03, 5.63, and 9.51, respectively (**Table 5**). For instance, a premenopausal patient who has HE4 of 38.4 pmol/L and 2 features on ultrasound findings would be scored as 3.21 (i.e., $0.04 \times 38.4 + 0 + 1.68 \times 1$); she is classified as medium risk of ovarian cancer. A postmenopausal woman who has HE4 of 75.8 pmol/L without ultrasound features would be scored as 3.85 (i.e., $0.04 \times 75.8 + 0.82 \times 1 + 0$), and thus will be classified as high risk of ovarian cancer.

A bootstrap technique with 200 replications was applied to assess internal validity of the HE4 model. The HE4 logit model

as in **Table 3** was constructed for each replicate data, and then risk prediction scores and probability of ovarian cancer were calculated. The estimated Somer'D correlation coefficients between observed and predicted values for the original and the bootstrap models were 0.787 and 0.771, respectively. The mean bias, a difference in observed versus predicted values, was only 1.97% (95% CI, 0.55 to 3.39). The C statistic of the derived and validated models were not much different, i.e., 0.893 and 0.886. The average difference (i.e., a degree of optimism) was 0.86% (95% CI, 0.24 to 1.49), indicating that the risk prediction score could well discriminate ovarian cancer from benign patients in both derived and validated data.

DISCUSSION

Using an existing dataset that was previously described in Chan et al. [19], we developed and evaluated risk prediction scores for ovarian cancer that use CA-125, HE4, ultrasound features, and menopausal status in logistic regression models. In this study, the risk prediction model containing HE4, menopausal status, and ultrasound findings exhibited the best performance in discriminating cancer from benign tumors with a C statistic of 0.893. This model performed better than the model containing CA-125, menopausal status, and ultrasound findings. Adding CA-125 to the HE4 model did not show an improvement in the C statistic. In addition, the HE4 model performed better than ROMA with an 8.8% improvement of reclassification of cancer and 15.9% improvement in the reclassification of benign tumors when evaluating the data by quartiles. The HE4 model was internally validated with calibration and discrimination biases of 1.97% and 0.86%, respectively.

When evaluating the simplified HE4 model (**Table 5**), 45 malignant samples (77.2%) were classified as medium or high risk and 233 benign samples (86%) were classified as very-low, low, or low-medium risk. When compared with ROMA using the cutoffs recommended by the manufacturer (7.4%

for premenopausal women and 25.3% for postmenopausal women), ROMA correctly classified 70.2% of the malignant samples as high risk and 91.9% of the benign samples as low risk and, indicating that the simplified HE4 model has better sensitivity over ROMA, but ROMA has higher specificity. Both the simplified HE4 model and ROMA demonstrated better specificity than CA-125 alone, where only 68.3% of the benign samples were identified as low risk.

It is important to triage women that have a pelvic mass into low or high risk for malignancy because many studies have shown that women with ovarian cancer that have surgery performed by a trained specialist at specialty centers have improved survival [2]. Differentiating a benign from a malignant pelvic mass can be especially difficult in premenopausal women because a pelvic mass is not uncommon, and is most likely going to be benign. Kim et al. [23] showed that referring all women with a pelvic mass to a gynecologic oncologist would lead to the most cost effective method of treatment because then all women with malignancy would be seen by the specialist, but in reality not all women have easy access to a specialist and there aren't enough specialists to handle that volume of patients. Algorithms such as RMI and ROMA provide information to the physician to help in deciding when to refer a woman to the specialist or perform the operation locally.

Most women with suspected ovarian cancer will have an ultrasound prior to surgery, and ultrasound prediction models to determine the risk that a pelvic mass is malignant can be very good, especially in the hands of an expert sonographer. Kaijser et al. [18] showed that using the LR2 risk prediction model developed by IOTA gave better diagnostic performance than ROMA with an AUC of 0.952. However, this was done by expert sonographers and it remains to be seen if this performance can be expected from a less-experienced technician. Karlsen et al. [17] suggests that a biomarker algorithm might be beneficial for centers that don't have expert sonographers. They found that ROMA and RMI performed equally well for distinguishing patients at a high risk of ovarian cancer, and suggested that using a strict biomarker algorithm may be less subjective. Stiekema et al. [24] attempted to develop an algorithm that included both HE4 and radiological features, but found that HE4 performed so well on its own at distinguishing between benign and malignant masses that ultrasound didn't provide any benefit. However, they did see an improvement in discrimination when intra-abdominal metastases as seen on CT scan were included in the analysis.

Our study showed that the new score using combination of HE4, ultrasound, and menopausal status may be a better predictor of malignancy than RMI, which combines CA-125,

ultrasound, and menopausal status, or ROMA, which combines CA-125, HE4, and menopausal status. Additional studies are needed to support these findings.

CONFLICT OF INTEREST

TCA and WS are the Lecture honoraria from Abbott Diagnostics and member of Abbott Scientific Advisory Board (South Asia). JS and BS are employees of Abbott Diagnostics. SH was the former Abbott Diagnostics employee. The other authors declare that they have no conflict of interest.

ACKNOWLEDGMENTS

This study was partially supported by Abbott Diagnostics. The authors acknowledge Dr Ammarin Thakkinstian, PhD from the Section of Epidemiology and Biostatistics, Faculty of Medicine, Ramathibodi Hospital for her contribution on the statistical analysis of this study.

REFERENCES

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin* 2009;59:225-49.
2. du Bois A, Rochon J, Pfisterer J, Hoskins WJ. Variations in institutional infrastructure, physician specialization and experience, and outcome in ovarian cancer: a systematic review. *Gynecol Oncol* 2009; 112:422-36.
3. Bristow RE, Chang J, Ziogas A, Anton-Culver H. Adherence to treatment guidelines for ovarian cancer as a measure of quality care. *Obstet Gynecol* 2013;121:1226-34.
4. 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;97:922-9.
5. Moore RG, McMeekin DS, Brown AK, DiSilvestro P, Miller MC, Allard WJ, 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;112:40-6.
6. Abraham J. OVA1 test for preoperative assessment of ovarian cancer. *Community Oncol* 2010;7:249-51.
7. 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;40:355-9.
8. Jacobs I, Bast RC Jr. The CA 125 tumour-associated antigen: a review of the literature. *Hum Reprod* 1989;4:1-12.
9. Escudero JM, Auge JM, Filella X, Torne A, Pahisa J, Molina R. Com-

- parison of serum human epididymis protein 4 with cancer antigen 125 as a tumor marker in patients with malignant and nonmalignant diseases. *Clin Chem* 2011;57:1534-44.
10. Gao L, Cheng HY, Dong L, Ye X, Liu YN, Chang XH, et al. The role of HE4 in ovarian cancer: inhibiting tumour cell proliferation and metastasis. *J Int Med Res* 2011;39:1645-60.
 11. Moore RG, Brown AK, Miller MC, Skates S, Allard WJ, Verch T, et al. The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecol Oncol* 2008;108:402-8.
 12. Moore RG, Miller MC, Disilvestro P, Landrum LM, Gajewski W, Ball JJ, et al. Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. *Obstet Gynecol* 2011;118:280-8.
 13. Kim YM, Whang DH, Park J, Kim SH, Lee SW, Park HA, et al. Evaluation of the accuracy of serum human epididymis protein 4 in combination with CA125 for detecting ovarian cancer: a prospective case-control study in a Korean population. *Clin Chem Lab Med* 2011;49:527-34.
 14. Ruggeri G, Bandiera E, Zanotti L, Belloli S, Ravaggi A, Romani C, et al. HE4 and epithelial ovarian cancer: comparison and clinical evaluation of two immunoassays and a combination algorithm. *Clin Chim Acta* 2011;412:1447-53.
 15. Bandiera E, Romani C, Specchia C, Zanotti L, Galli C, Ruggeri G, et al. Serum human epididymis protein 4 and risk for ovarian malignancy algorithm as new diagnostic and prognostic tools for epithelial ovarian cancer management. *Cancer Epidemiol Biomarkers Prev* 2011;20:2496-506.
 16. Moore RG, Jabre-Raughley M, Brown AK, Robison KM, Miller MC, Allard WJ, et al. Comparison of a novel multiple marker assay vs 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.
 17. Karlsen MA, Sandhu N, Hogdall C, Christensen IJ, Nedergaard L, Lundvall L, et al. Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. *Gynecol Oncol* 2012;127:379-83.
 18. Kaijser J, Van Gorp T, Van Hoorde K, Van Holsbeke C, Sayasneh A, Vergote I, 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;129:377-83.
 19. Chan KK, Chen CA, Nam JH, Ochiai K, Wilailak S, Choon AT, et al. The use of HE4 in the prediction of ovarian cancer in Asian women with a pelvic mass. *Gynecol Oncol* 2013;128:239-44.
 20. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157-72.
 21. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
 22. Schumacher M, Hollander N, Sauerbrei W. Resampling and cross-validation techniques: a tool to reduce bias caused by model building? *Stat Med* 1997;16:2813-27.
 23. Kim KH, Zsebik GN, Straughn JM Jr, Landen CN Jr. Management of complex pelvic masses using a multivariate index assay: a decision analysis. *Gynecol Oncol* 2012;126:364-8.
 24. Stiekema A, Lok CA, Kenter GG, van Driel WJ, Vincent AD, Korse CM. A predictive model combining human epididymal protein 4 and radiologic features for the diagnosis of ovarian cancer. *Gynecol Oncol* 2014;132:573-7.

