

## RESEARCH COMMUNICATION

# Ultrasound Operators' Confidence Influences Diagnosis of Ovarian Tumors - a Study in China

Huang Haiyan<sup>1</sup>, Deng Min<sup>2\*</sup>

### Abstract

**Aim:** To assess the effect of ultrasound operators' confidence in diagnosis of ovarian cancer, and the factors influencing diagnostic accuracy. **Methods:** Ultrasound images of selected ovarian cancers and controls were evaluated by 8 sinologists who were instructed to diagnose and classify lesions into benign, borderline or malignant, and we use structured questionnaire to investigate the level of confidence. We analyzed the accuracy of diagnosis, including sensitivity, specificity, positive and negative likelihood ratios and accuracy depending on the different levels of confidence. In addition, factors influencing diagnostic accuracy was assessed by logistic regression analysis. **Results:** A total of 426 cases were examined. The confidence score was significantly increased with the level of accuracy (test for trend,  $p < 0.05$ ). Borderline tumors were most difficult to diagnose, and had lower accuracy, sensitivity and specificity compared with benign and primary invasive tumors. Working experience was positively closely associated with diagnosis accuracy. Logistic regression analysis revealed working experience and confidence score to be positively related to the diagnostic accuracy (OR, 95% CI, 1.68, 1.15-3.97 for working experience; OR, 95% CI, 3.75, 1.67-6.98 for confidence score). **Conclusion:** Our study showed that level of confidence is positively associated with diagnostic performance, and the accuracy is greatly influenced by working experience and confidence score.

**Keywords:** Ultrasound operators - confidence and experience - ovarian tumor diagnosis - China

*Asian Pacific J Cancer Prev*, 12, 1275-1277

### Introduction

The overall mortality rate for ovarian tumor is 75%, but when diagnosed at stage I, 90% of patients can be cured if early diagnosis. Strategies for early detection require high sensitivity and extremely high specificity to attain a positive predictive value of at least 10% (Ren et al., 2010). Surgical removal is the standard treatment of an ovarian mass because of the difficulty in benign from malignant tumors prior to surgery (Tempany et al., 2000). The color Doppler ultrasonographic imaging is used to diagnose the ovarian cancer. The subjective evaluation by expert sinologists is superior to the use of scoring systems and mathematical models when classifying ovarian tumor into benign or malignant (Valentin et al., 1999; 2001; 2009). Previous study assessed the subjective evaluation by only two sonologists with about 97% sensitivity and 89% specificity (Timmerman et al., 1999). The confidence and experience of ultrasound operators, physical characteristics of patients and patient's ability to cooperate with examination could greatly influence the accuracy of diagnosis.

The confidence of ultrasound operators in determining diagnosis of ovarian tumor is usually determined by their personality, skill and experience, and the complex of

tumors. Previous study explored the association between confidence and diagnostic accuracy, and indicated the confidence of diagnosis influence the ultrasound diagnosis. But the study did not explore other factors which influenced the diagnosis accuracy, such as experience and complex of tumor. The confidence of operators only used three levels which could not better reflect the real situation. The aim of our study is to assess how confidence of ultrasound operators influences the diagnostic accuracy, and the influencing factors for diagnosis accuracy.

### Materials and Methods

#### *Study design and sample*

This was a prospective and observational study. 426 women who were histological diagnosed with ovarian tumor were selected Department of Ultrasonography in Guangdong Medical University from October 2008 and October 2010. The cases included a mix of representative examples of benign, borderline and invasive malignant ovarian tumors.

#### *Measurements*

Color Doppler images were used for all the cases, and we recorded the reports which contained information on

<sup>1</sup>Ultrasonography Department, The Affiliated Hospital of Guangdong Medical College, Zhanjiang, <sup>2</sup>Department of Imaging and Interventional Radiology, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong, China \*For correspondence: dengmin56@yahoo.com.cn

**Table 1. Test Performance Characteristics by Level of Confidence, Histology and Working Year**

	N(%)	Accuracy	Sensitivity	Specificity	LR+	LR-
<b>Confidence</b>						
<4	16 (3.8)	31.3	14.3(0.4-58.0)	44.4(13.7-78.8)	0.26(0.04-1.73)	1.93(0.88-4.25)
5	37 (8.7)	70.3	66.7(22.3-95.7)	71.0(52.0-85.8)	2.29(1.04-5.06)	0.47(0.15-1.49)
6	58 (13.6)	81.0	72.7(39.0-94.0)	82.9(69.2-92.4)	4.27(2.06-8.84)	0.33(0.12-0.87)
7	77 (18.1)	88.3	86.7(63.6-98.5)	88.7(78.1-95.3)	7.68(3.72-15.9)	0.15(0.04-0.54)
8	82 (19.2)	90.2	88.2(63.6-98.5)	90.8(81.0-96.5)	9.56(4.37-20.9)	0.13(0.04-0.48)
9	79 (18.5)	96.2	94.7(74.0-99.9)	96.7(88.5-99.6)	28.4(7.3-111.5)	0.05(0.01-0.37)
10	77(18.1)	98.7	100(85.2-100)	98.1(90.1-100)	35.9(7.4-174.3)	0.02(0.001-0.3)
Total	426	87.3(72.5-93.6)	83.7(74.8-90.4)	88.4(84.4-91.7)	6.01(2.61-13.9)	0.22(0.06-0.79)
>6	373 (87.6)	91.7(82.5-98.7)	90.6(82.3-95.8)	91.7(87.9-94.6)	10.5(4.90-22.1)	0.14(0.06-0.32)
>8	238 (55.9)	94.9(85.2-99.5)	94.9(85.9-98.9)	95.0(90.7-97.7)	18.1(6.80-48.4)	0.08(0.03-0.22)
<b>Histology</b>						
Benign	217 (51%)	88.5(79.7-96.9)	87.1(81.9-91.3)	90.0(85.1-93.7)	8.67(5.76-13.05)	0.14(0.10-0.20)
Borderline	111 (26%)	72.5(61.4-83.2)	64.9(55.2-73.7)	75.2(70.1-79.9)	2.62(2.06-3.32)	0.47(0.36-0.61)
Invasive	98 (23%)	87.3(72.5-93.6)	83.7(74.8-90.4)	88.4(84.4-91.7)	6.01(2.61-13.85)	0.22(0.06-0.79)
<b>Working years</b>						
<5	2 (25.0%)	88.5(79.7-96.9)	87.1(81.9-91.3)	90.0(85.1-93.7)	8.67(5.76-13.05)	0.14(0.10-0.20)
5-10	3 (37.5%)	72.5(61.4-83.2)	64.9(55.2-73.7)	75.2(70.1-79.9)	2.62(2.06-3.320)	0.47(0.36-0.61)
>10	3 (37.5%)	87.3(72.5-93.6)	83.7(74.8-90.4)	88.4(84.4-91.7)	6.01(2.61-13.85)	0.22(0.06-0.79)

(95%CI)

the color score by sonologist into an enclosed envelope. All the masses were histological diagnosed as the gold standard in determining the correctness of the ultrasound determination. We gave the anonymous images of masses to 8 observers. The score of image and histological results was blinded to all the observers. But we gave them the clinical information of patients, such as clinic symptoms, indication for the scan, family history and personal information. For each mass, we used a visual scale from 1 to 10 to answer the confidence in diagnosis, and the question was ‘How confidence are you of your findings?’, and evaluated whether they were influenced by the former diagnosis with the question ‘Whether this diagnosis were influenced by the former diagnosis’. Then the observers were divided the mass into benign, borderline or invasive malignant ovarian tumors, and chose the histological diagnosis of the mass. We recorded the basic information of observers, including working year and formal diagnostic study interpretation.

**Data analysis**

The overall accuracy, sensitivity, specificity, and likelihood ratios for a positive (LR+) and negative (LR-) test were calculated for all examinations together and following stratification into different histological types and working experience. Logistic regression was performed modeling the following variables: working year, formal diagnostic study interpretation, tumor types and operator confidence. Additionally, the spearman correlation was used to test the association between the correlation of tumor type and confidence value. StatsDirect™ was used for all statistical calculations.

**Results**

A total of 426 women diagnosed with ovarian tumors were collected. All the mass images of 426 women were assessed by the 8 examiners. Among these 426 women, 51%(217 masses) of them were benign tumors, 26%(111

masses) of them were borderline ovarian tumors and 23%(98 masses) of them were primary invasive ovarian tumors. The mean age of women among benign, borderline and invasive ovarian tumors were 43±8 years, 45±11 years and 38±9 years, respectively.

The test performance by confidence, histological type and working year score were showed in Table 1. The total accuracy, sensitivity, specificity, LR+ and LR- were 87.3%, 83.7%, 88.4%, 6.01 and 0.22, respectively. The confidence score was significantly increased with the raising of accuracy (test for trend, p<0.05). When the score was lower than 4, the accuracy was as low as 31.3% to diagnose ovarian tumor. While, when the score was higher than 8, the accuracy could be as high as 95%. The borderline tumor was seemed lowest accuracy, sensitivity and specificity compared with benign and primary invasive tumors, which indicated the borderline tumor was difficult to diagnose. Working experience was closely associated with diagnosis accuracy, and the long working year was related to the high accuracy (p<0.05).

We further analyzed the relationship between the confidence score, histology type and accuracy. The confidence score was significantly positive related to the diagnostic accuracy, and the contingency coefficient showed the confidence score was closely associated with benign and primary invasive tumor (Benign and primary invasive contingency coefficient: 0.73, p<0.001 and 0.67, p<0.001, respectively), and moderated associated with borderline tumor(contingency coefficient: 0.41, p<0.001).

Logistic regression analysis revealed working experience and confidence score was positively related to the diagnostic accuracy when adjusting formal diagnosis and tumor type (OR, 95%CI, 1.68, 1.15-3.97 for working experience; OR, 95%CI, 3.75, 1.67-6.98 for confidence score).

**Discussion**

Our study indicated that a clear association between

the level of confidence and the diagnostic performance of the operator in determining the type of ovarian tumors. The diagnostic accuracy of the diagnosis declined as the confidence in making the diagnosis decreased. The accuracy showed great difference in diagnosis of benign, borderline or primary invasive malignant ovarian tumors, and the borderline tumor is hardest to diagnose and determine. Besides, the accuracy is modified by the working year and confidence score.

The knowledge and experience of operator is usually associated with the diagnostic accuracy and certainty. The confidence is extremely difficult to assess, as it is influenced by complex factors which could interact and influence the level of confidence. The influencing factors include the amount of training or working years, time allocated to diagnosis, memory and emotions, etc. (Baranski and Petrusic, 1998; Rohrbaugh and Shanteau, 1999). The results of our study indicated the diagnostic accuracy is affected by the operator working years and confidence scores, and the confidence score greatly influence the accuracy. This could be explained that the confidence score is represented the training, working experience, memory and emotions, etc.

The diagnostic accuracy showed the lowest in the case of borderline ovarian tumors, which were most often misclassified as benign tumors. This may be because certain types of benign and borderline ovarian tumor usually share similar morphological features, which make it difficult to diagnose between the two types of tumor. In our study, the confidence score is high associated with benign and primary invasive tumors and the moderated associated with borderline tumor. This indicated the difficult of morphological features influenced the confidence score and accuracy. If the borderline ovarian tumor were misclassified into other type of tumor, it may induce overtreatment or miss diagnosis to treatment delayed.

The ultrasound examination is ideally to differentiate accurately between benign, borderline and invasive malignant ovarian tumors. But in practice, most ultrasound operators classify ovarian tumor as benign or malignant and only a few operators make diagnosis of borderline ovarian tumor. Our results indicated the scores of diagnosing borderline tumor is the lowest, which suggests the morphological features of the borderline ovarian tumors is hard to understand, and the accuracy of ultrasound diagnosis would be improved if the lesions were classified as invasive or non-invasive instead. Therefore, many borderline ovarian tumors receive similar treatment to benign lesions with a risk that their content would sometimes be spilled during minimally invasive surgery, and the misdiagnosis would have an adverse effect on long-term outcomes and survival rates (Fauvet et al., 2005).

There were several limitations in our study. Firstly, the subjects in our study is the tumor patients with a high proportion of tumors that were not obvious benign or malignant, and it is hard to distinguish them. But in actually situation, the accuracy of diagnosis would be decreased in a non-selected tumor population. Therefore, there would be selection bias in this analysis. Secondly, the

confidence score was assessed by a self-assessment visual scale and several simple questions. This self-assessment is subjective and may be earlier affected by the individuals' emotions and attitude to their own confidence, and this may be induced measurement bias in our study. Therefore, it needs a more objective way to assess the confidence of operator.

In conclusion, our study has shown that level of confidence is positive association with the diagnostic performance, and the accuracy is influenced by the histological type of ovarian tumor. Borderline ovarian tumor is the main source of diagnostic uncertainty. If the diagnosis is probable, the morphological description of the tumor should be accompanied by a list of possible histological diagnosis to improve the diagnostic accuracy, because the ultrasound diagnosis is influenced by the level of confidence and working experience of operator. It is reasonable that the ultrasound report should add the report about the ultrasound examination is not informative enough to make diagnosis, and this is the responsibility of the ultrasound operator for the patients' health.

## References

- Baranski JV, Petrusic WM (1998). Probing the locus of confidence judgments: experiments on the time to determine confidence. *J Exp Psychol Hum Percept Perform*, **24**, 929-45.
- Fauvet R, Boccara J, Dufournet C, et al (2005). Laparoscopic management of borderline ovarian tumors: results of a French multicenter study. *Ann Oncol*, **16**, 403-10.
- Ren J, Cai H, Zhang X, et al (2010). Tumor markers for early detection of ovarian cancer. *Expert Review of Molecular Diagnostics*, **10**, 787-98.
- Rohrbaugh C, Shanteau J (1999). Context, process and experience: research on applied judgment and decision making. New York: John Wiley.
- Tempany CM, Zou KH, Silverman SG, et al (2000). Staging of advanced ovarian cancer: comparison of imaging modalities: report from the Radiological Diagnostic Oncology Group. *Radiology*, **215**, 761-767.
- Timmerman D, Schwarzler P, Collins WP, et al (1999). Subjective assessment of adnexal masses with the use of ultrasonography: an analysis of interobserver variability and experience. *Ultrasound Obstet Gynecol*, **13**, 11-16.
- Valentin L (1999). Prospective cross-validation of Doppler ultrasound examination and grayscale ultrasound imaging for discrimination of benign and malignant pelvic masses. *Ultrasound Obstet Gynecol*, **14**, 273-283.
- Valentin L, Jurkovic D, Van Calster B, et al (2009). Adding a single CA-125 measurement to ultrasound performed by an experienced examiner does not improve preoperative discrimination between benign and malignant adnexal masses. A prospective international multicentre study of 809 patients. *Ultrasound Obstet Gynecol*, **34**, 345-354.
- Valentin L, Hagen B, Tingulstad S, et al (2001). Comparison of 'pattern recognition' and logistic regression models for discrimination between benign and malignant pelvic masses: a prospective cross validation. *Ultrasound Obstet Gynecol*, **18**, 357-365.