OC201
The interobserver agreement between expert ultrasound operators when differentiating between adnexal tumors using ultrasound pattern recognition.

J. Yazbek[1], T. K. Holland[1], C. Van Holsbeke[2], A. Testa[3], L. Valentin[4], D. Jurkovic[1], D. Timmerman[2]

[1]*Early Pregnancy and Gynaecology Assessment Unit, King's College Hospital, London, United Kingdom,* [2]*University Hospitals Leuven, Belgium,* [3]*Istituto di Clinica Ostetrica e Ginecologica, Università Cattolica del Sacro Cuore, Italy,* [4]*Department of Obstetrics and Gynecology, University Hospital Malmö, Sweden*

Objectives: To assess interobserver agreement and diagnostic accuracy of expert ultrasound operators using pattern recognition to assess the nature of adnexal masses.
Methods: Static B-mode preoperative ultrasound images containing gray-scale and color Doppler information of a selected mix of difficult adnexal masses of 166 patients were examined independently by three expert ultrasound operators (A, B and C). They all had access to relevant clinical information, but none of the experts performed the original real-time scans. The experts were asked to classify tumors as benign or malignant and to subclassify tumors into one of 11 histological subgroups. The diagnostic performance of each expert was compared with the histological diagnosis of the respective specimen. Each operator's diagnoses were compared with the other two operators' diagnoses using Cohen's kappa coefficient.
Results: There were 36 invasive malignancies, 34 borderline ovarian tumors (BOT) and 96 benign tumors. The sensitivity and specificity with regard to malignancy (BOT and invasive) were: 86% (95% CI, 75–93) and 91% (95% CI, 83–96) for Expert A, 86% (95% CI, 75–93) and 80% (95% CI, 70–87) for Expert B, and 80% (95% CI, 68–89) and 85% (95% CI, 77–92) for Expert C. There was a good interobserver agreement when diagnosing ovarian tumors as benign, BOT and invasive (kappa = 0.69, 95% CI, 0.59–0.79 comparing Experts A and B; kappa = 0.71, 95% CI, 0.62–0.82, comparing Experts A and C; and kappa = 0.75, 95% CI, 0.66–0.84, comparing Experts B and C) and when predicting the histological subtypes (kappa = 0.63, 95% CI, 0.55–0.71, comparing Experts A and B; kappa = 0.67, 95% CI, 0.59–0.75, comparing Experts A and C; and kappa = 0.66, 95% CI, 0.58–0.76, comparing Experts B and C).
Conclusions: Expert ultrasound operators have a good level of agreement when differentiating between 'difficult' ovarian tumors using the ultrasound pattern recognition method. This shows that the high accuracy of this method is reproducible.

OC202
Real-time ultrasound versus static image evaluation of adnexal masses

C. Van Holsbeke[1], J. Yazbek[2], A. Daemen[3], T. Holland[2], A. Testa[4], L. Valentin[5], D. Timmerman[1], D. Jurkovic[2]

[1]*University Hospitals Leuven, Belgium,* [2]*Early Pregnancy and Gynaecology Assessment Unit, King's College Hospital, London, United Kingdom,* [3]*Department of Electrical Engineering, ESAT-SCD, Katholieke Universiteit Leuven, Belgium,* [4]*Istituto di Clinica Ostetrica e Ginecologica, Università Cattolica del Sacro Cuore, Roma, Italy,* [5]*Department of Obstetrics and Gynecology, University Hospital Malmö, Sweden*

Objectives: Expert sonologists may use their subjective impression ('pattern recognition') to achieve a high accuracy in differentiation between benign or malignant adnexal masses. The aim of this study was to establish whether the ability to achieve a correct diagnosis differs between an expert performing real-time examination in comparison with experts making a diagnosis by evaluating static ultrasound images stored in the database.
Methods: A database was searched to identify 166 patients with an ultrasound diagnosis of adnexal tumors, who all subsequently underwent surgery. In all cases expert sonologists had access to all relevant clinical information. A single expert performed real-time examinations, whilst the stored static database images, containing gray-scale and color Doppler information on the adnexal masses, were independently examined by three other expert sonologists. All four sonologists had to classify the mass as benign or malignant. They also had to define the level of confidence with which they made the diagnosis. In cases of disagreement between the experts reviewing the images, the histological diagnosis made by two of the three examiners was taken as being representative of the particular case. The gold standard was the final histology after surgery. The McNemar test was used for statistical analysis.
Results: In 134/166 (80.7%) patients both the real-time examiner and the picture experts achieved a correct diagnosis of a benign or malignant adnexal tumor. In 14/166 (8.4%) cases they both made an incorrect diagnosis. In the remaining 18/166 (10.8%) cases where the experts disagreed, the real-time sonologist made a correct diagnosis in 14 cases and the image reviewing experts in four cases ($P = 0.0184$).
Conclusions: An expert performing a real-time ultrasound examination has a significantly better chance of achieving a correct diagnosis of a benign or malignant adnexal tumor compared with experts reviewing static ultrasound images.

OC203
How accurate is ultrasound pattern recognition at predicting the histological diagnosis of an ovarian mass?

J. Yazbek[1], C. Van Holsbeke[2], A. Daemen[3], T. K. Holland[1], A. Testa[4], L. Valentin[5], D. Timmerman[2], D. Jurkovic[1]

[1]*Early Pregnancy and Gynaecology Assessment Unit, King's College Hospital, London, United Kingdom,* [2]*University Hospitals Leuven, Belgium,* [3]*Department of Electrical Engineering, ESAT-SCD, Katholieke Universiteit Leuven, Belgium,* [4]*Istituto di Clinica Ostetrica e Ginecologica, Università Cattolica del Sacro Cuore, Roma, Italy,* [5]*Department of Obstetrics and Gynecology, University Hospital Malmö, Sweden*

Objectives: To assess the accuracy of pattern recognition for the histological diagnosis of an adnexal mass, when the examinations are performed by ultrasound experts of similar experience.
Methods: Static B-mode preoperative ultrasound images, containing gray-scale and color Doppler information on the adnexal masses of 166 patients were examined independently by three expert sonologists. They all had access to relevant clinical information, but none of the experts performed the original real-time scans. The expert sonologists were asked to classify tumors into one of 11 histological groups. They were also asked to indicate the degree of confidence with which they made the diagnosis. In cases of disagreement between the experts reviewing the images, the histological diagnosis made by two of the three examiners was taken as the representative of the particular case. The gold standard was the final histology.
Results: As a group the experts reached an accuracy of 83.13% in classifying the adnexal mass as benign or malignant. In six patients all three examiners gave a different histological diagnosis and these cases were excluded from further analysis. The sensitivity and specificity for the different histologies were: 91.43% (32/35) and 97.60% (122/125) for dermoid cysts; 66.67% (22/33) and 90.55% (115/127) for cystadenoma (fibroma); 93.33% (14/15) and 99.31% (144/145) for endometrioma; 68.75% (22/32) and 90.63% (116/128) for borderline ovarian tumors (BOT); 42.86% (6/14) and 95.89% (140/146) for gastrointestinal BOTs; 88.89% (16/18) and 95.77% (136/142) for serous BOTs; 88.00% (22/25) and 99.26%