Evaluation of the Four-dimensional “Spatiotemporal Image Correlation” Technology with High-definition Color Doppler as Third Step for Preoperative Differential Diagnosis of Ovarian Tumors: A Prospective Study

Juan L Alcázar et al

ABSTRACT

Aim: The purpose of this study is to analyze whether the addition of volumetric pulsatility indices (vPI) derived from spatiotemporal image correlation (STIC) assessment could improve the diagnostic performance of conventional two-dimensional (2D) ultrasound for discriminating benign from malignant adnexal tumors used as a third sequential step.

Study design: Ours was an observational prospective cohort study between September 2014 to September 2016. Women diagnosed as having a persistent ovarian tumor evaluated at our center were included in this study. All women underwent transvaginal ultrasonography before surgery to assess the ultrasonographic characteristics of the lesion by gray-scale morphological analysis classifying the lesion as benign or suspicious based on “pattern recognition”. In benign lesions, no further test was conducted. For those tumors regarded as suspicious, assessment of the vascularization was performed by using conventional 2D high-definition flow Doppler (2D-HDF), classifying the lesion as suspicious if vascularization within solid component was detected. In those lesions with intratumoral vascularization, four-dimensional (4D)-STIC technology with HDF was applied for calculating vPI. We used two reference standards: Definitive histology when the mass was surgically removed and in case the mass was not removed, those with spontaneous resolution or persistent without change and follow-up longer than 12 months were considered as benign. We compared diagnostic performance of conventional approach (pattern recognition followed by 2D-HDF) vs pattern recognition followed by 2D-HDF followed by 4D-STIC.

Results: A total of 214 patients were included; 175 had a benign tumor and 39 a malignant tumor. We observed that vPI was significantly lower in malignant masses as compared with benign masses when 4D-STIC was used. Sensitivity and specificity of B-mode + 2D-HDF were 0.95 (95% confidence interval (CI): 0.83–0.98) and 0.92 (95% CI: 0.87–0.95) respectively. Sensitivity and specificity of B-mode + 2D-HDF + 4D-STIC were 0.95 (95% CI: 0.83–0.98) and 0.94 (95% CI: 0.90–0.96) respectively. No significant differences by adding the STIC were observed (McNemar test p > 0.05).

Conclusion: The 4D-STIC does not improve the diagnostic performance of conventional ultrasound in the differential diagnosis of ovarian tumors as compared with conventional assessment using 2D ultrasound. However, a low vPI increases significantly the probability of malignancy in questionable masses.

Keywords: Adnexal masses, Power Doppler, Spatiotemporal image correlation, Two-dimensional, Ultrasound.

INTRODUCTION

Transvaginal ultrasound has been shown as an excellent diagnostic tool for discriminating benign from malignant adnexal masses when performed by expert examiners. Most ovarian cancers appear as solid or multilocular-solid masses with vascularization as detected by color or power Doppler within those solid components and many times with signs of intra-abdominal spread, such as ascites or frank carcinomatosis. On the contrary, adnexal cysts with a typical pattern of a benign lesion, such as simple cyst, endometrioma, dermoid, hydrosalpinx, and peritoneal cyst, are rarely malignant lesions.

The finding of solid tissue within a cyst or the presence of a solid lesion raises the suspicion of a possible malignancy. However, many benign masses exhibit this ultrasound appearance. Some studies have found that this kind of adnexal masses are quite difficult to categorize, even by expert examiners, when no other signs of malignancy like ascites or carcinomatosis are present. The addition of 2D color/power Doppler may reduce significantly the false-positive rate of B-mode ultrasound in
these cases. In spite of this, about 15% of benign lesions exhibit vascularized solid components. Since the pioneering report from Kurjak et al., dimensional power Doppler angiography (3D-PDA) has been proposed as a method to improve our ability to distinguish benign from malignant adnexal masses. Some studies have shown that the so-called virtual vascular sampling and 3D vascular indices calculation could be useful in some selected adnexal masses. However, other studies have challenged this.

We performed a prospective study using static 3D-PDA as a third step in the differential diagnosis of adnexal masses and we found out that the addition of 3D-PDA did not provide additional information to conventional 2D B-mode/color Doppler ultrasound and did not improve diagnostic performance.

The advent of 4D angiography (4D visualization combined with Doppler options) together with the option called STIC allowed detailed estimation of blood flow changes in the tissue under investigation within one cardiac cycle. It paved the way for using 4D angiography in differentiating pathological tissues. Using this technology, a new volumetric vascular index has been developed, the so-called volumetric pulsatility index.

In this study, we aimed to assess whether the addition of vPI derived from STIC assessment could improve the diagnostic performance of conventional 2D ultrasound for discriminating benign from malignant adnexal tumors used as a third sequential step.

**MATERIALS AND METHODS**

This is a cohort prospective observational study performed at a single tertiary care university center between September 2014 to September 2016, comprising a series of consecutive women.

All women gave oral consent and institutional review board approval (CUN-37/2014) was obtained before starting the study.

Eligible patients were all women with a diagnosis of a persistent adnexal masses evaluated at our center during the study period. Inclusion criteria were patients aged between 18 and 80 years who accepted transvaginal or transrectal ultrasound and who provided informed consent. Pregnant women, those who were managed expectantly and have less than 1-year follow-up, or those who underwent surgery out of our center were excluded.

All included patients underwent transvaginal or transrectal ultrasound by staff gynecologists (MA, ARZ, BO, and TE) or one expert gynecologist with special interest on gynecological ultrasound (JLA). Transabdominal ultrasound was additionally performed in cases of large masses or suspicious masses for assessing intra-abdominal spread.

All examinations were performed using a Voluson E8 machine (GE Health care, Milwaukee, USA).

Ultrasound scanning protocol was as follows: Initially, all masses were evaluated by gray-scale ultrasound according to pattern recognition approach and using International Ovarian Tumor Analysis (IOTA) terms and definitions. Those masses with a typical appearance of a benign cyst, such as endometrioma, mature teratoma, simple/serous cyst, hydrosalpinx, paraovarian cyst, hemorrhagic cyst, mucinous cyst, or peritoneal cyst were considered as low risk for malignancy and did not undergo further assessment by ultrasound.

Those masses that did not show ascites and/or signs of intra-abdominal carcinomatosis that could not be classified as certainly benign according to pattern recognition and showed solid components (thick papillary projections, irregular internal wall, or solid areas) or the tumor that was solid were evaluated by 2D-HDF angiography to determine the location and amount of color signals (IOTA color score) within the mass and specifically within the solid components. High-definition flow Doppler settings were manually adjusted in each case in order to achieve maximum sensitivity without noise for detecting color signals. If no color signals (color score 1) were detected within solid components, the mass was considered as low risk. However, if color signals were detected within the solid components (color score of 2, 3, or 4), the mass was considered as high risk. These masses were then evaluated using 4D-STIC technology.

Those cases with ascites and/or sonographic signs for carcinomatosis detected by B-mode were considered as high risk for malignancy. In spite of evidence of intra-abdominal spread of disease, these masses were also evaluated by 4D-STIC technology.

For 4D-STIC assessment, we used a standardized approach. We first selected the area with the highest vascularization as depicted by high-definition flow and closer than 5 cm to the vaginal probe, then STIC was activated and a 4D-STIC volume was acquired while the patient remained as still as possible. The HDF Doppler settings were set as follows: HDF window angle, 70°; zoom, 1.4; depth of HDF window, 40 mm; pulse repetition frequency, 0.6 kHz; gain, 0.2; frequency, “mid”; HDF quality, “normal”; wall motion filter, “low 1”; smooth, 3/5; ensemble, 10; flow resolution, “mid2”; line density, 9; balance, >180; HD map, 1. The HDF refreshing frequency resulting from these settings was 15 Hz. The STIC sweep angle was 40° and the acquisition time 15 seconds.

The 4D-STIC offline analyses were carried out by the same examiner (JLA) using 4D-View version 10.4 software (GE Medical Systems, Zipf, Austria) on a personal computer. The first step in calculations for every patient was to analyze the STIC loop: After magnifying...
each image and bringing the region with the largest content of color-coded signal in the STIC representation to the center of the screen (in the y-axis), each of the 3D HDF volumes forming the loop was saved. All 3D frames from a reconstructed cardiac cycle were then analyzed consecutively, the center of a 1 cm$^3$ virtual sphere being placed in the center of the area with highest color content. The sphere volume size of 1 cm$^3$ (tolerance limit $\leq 0.05$ cm$^3$) was set by the operator using the semiautomatic software. The HDF histograms were then calculated automatically for each 3D volume of the STIC reconstructed cardiac cycle, providing VI value. We then selected the highest ($V_{I_{\text{sys}}}$) and the lowest ($V_{I_{\text{diast}}}$) VI values obtained, assuming that they represent systole and diastole respectively (Figs 1 and 2). We also calculated VI mean for one cardiac cycle, based on all volumes in the STIC sequence. Using these VI values, we calculated the vPI according to the following formula: 

$$vPI = (\text{systolic VI} - \text{diastolic VI})/\text{mean VI}.$$ 

Mean time spent for performing this was 11.3 minutes, ranging from 9.1 to 13.6 minutes.

Following the ultrasound examination, patients ultimately included were managed expectantly (asymptomatic women with a diagnosis of a presumed functional or benign lesion) or were submitted to surgery (patients with a suspicious mass, patients presenting complaints or fertility problems or because of patient’s decision). We used two reference standards: One, definitive histological diagnosis when the mass was surgically removed; and two, those masses that resolved spontaneously or had no change after 1-year follow-up were considered as benign for analytical purposes. Tumors surgically removed were classified according to the World Health Organization criteria$^{25}$ and ovarian cancers were staged according to International Federation of Gynecology and Obstetrics criteria.$^{26}$ Ovarian borderline tumors (BOT) were considered as malignant for analysis.

Clinical decisions were taken according to clinical data and ultrasound data from gray-scale and 2D-HDF ultrasound. Clinicians, surgeons, and pathologists were not aware about 4D-STIC data. Pathologists were aware of 2D ultrasound findings.

Figs 1A and B: The 4D-STIC assessment of a solid area within a cystic mass during systole. Note vascular map with the solid area (A). Vascularization index was 42.509% (B)

Figs 2A and B: Same case as in Figure 1 but 4D-STIC assessment during diastole. Note that vascular map within the solid area is less rich than in Figure 1 (A). Vascularization index was 39.284% (B)
Qualitative variables were expressed as absolute number and percentage. The Kolmogorov–Smirnov test was used to check normal distribution of data of quantitative variables. One-way analysis of variance or Mann–Whitney U-tests were used to compare continuous variables.

Receiver-operating characteristic curves were plotted to determine the best vPI cut-off for differentiating benign from malignant masses. This was selected according to the best combination of high sensitivity and low false-positive rate.

Sensitivity, specificity, positive likelihood ratio and negative likelihood ratio (with 95% CI) for the strategy based on gray-scale B-mode and 2D-HDF as well as the strategy based on gray-scale B-mode and 2D-HDF followed by 4D-STIC were calculated. Sensitivity and specificity were compared using McNemar’s test.

Receiver operating characteristic (ROC) curve was plotted to determine the optimal vPI cut-off to be used, considering to keep a high sensitivity with acceptable specificity. The 4D-STIC result was considered as high risk for malignancy with vPI was below the cut-off as determined by ROC curve analysis. Therefore, tumor classification based on 4D-STIC result was retrospectively estimated.

Since there are no data reported regarding the use of 4D-STIC in the differential diagnosis of adnexal masses, we could not calculate the sample size for we were not able to determine the minimal difference clinically relevant in terms of sensitivity and specificity when comparing gray-scale B-mode and 2D-HDF and gray-scale B-mode and 2D-HDF followed by 4D-STIC.

A p-value <0.05 was considered as statistically significant for all comparisons. The STROBE and STARD guidelines were followed for reporting this study.27,28

RESULTS

During the study period, 270 women were eligible for the study. Fifty-six women were excluded because of pregnancy at the time of diagnosis (n = 5), surgery was performed out of our center (n = 15), and 4D-STIC technology was not applied because of the area of interest within the mass was farther than 5 cm from the transvaginal probe (n = 4) or expectant management with less than 1-year follow-up (n = 32). Thus, 214 women were ultimately included.

Patients’ mean age was 46.4 years [standard deviation (SD): 13.8], ranging from 18 to 77 years; 133 (62.1%) were premenopausal and 81 (37.9%) were postmenopausal. This was selected according to the best combination of high sensitivity and low false-positive rate.

Patients’ mean age was 46.4 years [standard deviation (SD): 13.8], ranging from 18 to 77 years; 133 (62.1%) were premenopausal and 81 (37.9%) were postmenopausal. Clinically, most patients were asymptomatic (n = 131, 61.2%), 66 (30.8%) women complained pelvic or low back abdominal pain, 11 (5.1%) women complained fertility problems, and 6 (2.8%) women complained abdominal symptoms (abdominal discomfort, swelling, bloating, constipation, diarrhea, or increase in abdominal size).

A total of 182 (85%) women were managed surgically, all of them within 4 weeks after ultrasound evaluation. After tumor removal, 143 (78.6%) masses were proven to be benign and 39 (21.4%) were malignant (Table 1). Thirty-two women were managed expectantly. In 6 cases, the lesion resolved spontaneously and in 26 cases, the lesion remained unchanged after 1-year follow-up. All of them were considered as benign.

Among primary ovarian malignancies, tumor stage was as follows: Stage I in 16 cases (including 3 BOTs), stage II in 3 cases, and stage III in 13 cases.

Patients with benign tumors were younger (mean age: 45.2 years, SD: 13.6) than women with malignant tumors (mean age: 51.9 years, SD: 13.9) (p = 0.006).

Patients’ complaints according to reference standard are shown in Table 2.

The flow diagram of the patients is shown in Flow Chart 1. After gray-scale ultrasound and 2D-HDF, 163 (76.2%) masses were considered as low risk (146 benign pattern recognition and 17 with no color flow detected as for malignancy) and 51 (23.8%) as high risk for malignancy. Among those considered as low risk, reference standard was benign in 161 cases, and 2 cases were malignant. Among those 51 cases considered as high risk, 37 were confirmed as malignancies and 14 tumors were benign.

Table 1: Histologic diagnosis of tumors removed surgically (n = 182)

<table>
<thead>
<tr>
<th>Histology</th>
<th>n</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometrioma</td>
<td>51</td>
<td>28.0</td>
</tr>
<tr>
<td>Serous cyst</td>
<td>29</td>
<td>15.9</td>
</tr>
<tr>
<td>Dermoid cyst</td>
<td>22</td>
<td>12.1</td>
</tr>
<tr>
<td>Fibroma</td>
<td>13</td>
<td>7.1</td>
</tr>
<tr>
<td>Mucinous cyst</td>
<td>9</td>
<td>4.9</td>
</tr>
<tr>
<td>Cystadenofibroma</td>
<td>5</td>
<td>2.7</td>
</tr>
<tr>
<td>Tubo-ovarian abscess</td>
<td>4</td>
<td>2.2</td>
</tr>
<tr>
<td>Paraovarian cyst</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>Other benign*</td>
<td>5</td>
<td>2.7</td>
</tr>
<tr>
<td>BOT</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>Primary invasive ovarian cancer</td>
<td>29</td>
<td>15.9</td>
</tr>
<tr>
<td>Metastatic tumor</td>
<td>6</td>
<td>3.3</td>
</tr>
<tr>
<td>Other malignancy†</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

BOT: Borderline ovarian tumor; *Hemorrhagic cyst, Brenner tumor, Peritoneal cyst, Retroperitoneal lipoma, pedunculated adenomyoma; †One gastrointestinal stromal tumor

Table 2: Patients’ complaints according to reference standard

<table>
<thead>
<tr>
<th>Complaint</th>
<th>Benign tumors</th>
<th>Malignant tumors</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymptomatic</td>
<td>108</td>
<td>23</td>
<td>131</td>
</tr>
<tr>
<td>Pelvic/lower back pain</td>
<td>55</td>
<td>11</td>
<td>66</td>
</tr>
<tr>
<td>Fertility problems</td>
<td>11</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Abdominal swelling</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>175</td>
<td>39</td>
<td>214</td>
</tr>
</tbody>
</table>
The 4D-STIC was, therefore, performed in 51 women. Median vPI was significantly lower in malignant tumors (0.043, IQR: 0.07) as compared with benign tumors (0.187, IQR: 0.23) (p < 0.001). The best cut-off for discriminating benign from malignant tumors was a vPI value ≤0.31 (area under the curve: 0.834, 95% CI: 0.708–0.959).

After adding 4D-STIC, and considering the cut-off above mentioned, 167 masses were considered as benign (4 cases previously considered as high risk were now considered as low risk) and 47 as malignant; 165 out of 167 cases considered as low risk were benign and 2 were malignant. Thirty-seven out of 47 cases considered as high risk were malignant and 10 were benign.

Diagnostic performance of each strategy is shown in Table 3. We did not observe statistical differences (McNemar test, p = 0.125).

Ethical approval: All procedures performed in studies involving human participants were in accordance with

Table 3: Diagnostic performance of both approaches

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>LR+</th>
<th>LR−</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-mode + 2D Doppler</td>
<td>95%</td>
<td>92%</td>
<td>11.8</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>95% CI: 83–98%</td>
<td>95% CI: 87–95%</td>
<td>95% CI: 7.1–19.7</td>
<td>95% CI: 0.01–0.21</td>
</tr>
<tr>
<td>B-mode + 2D Doppler + 4D-STIC</td>
<td>95%</td>
<td>94%</td>
<td>16.6</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>95% CI: 83–98%</td>
<td>95% CI: 90–97%</td>
<td>95% CI: 7.6–54.0</td>
<td>95% CI: 0.01–0.21</td>
</tr>
</tbody>
</table>

LR: Likelihood ratio
the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Institutional review board approval (CUN-37/2014) was obtained before starting the study.

Informed consent: Informed consent was obtained from all individual participants included in the study.

**DISCUSSION**

In the present study, we have assessed the potential role of 4D-STIC technology as adjuvant tool to conventional 2D gray-scale and color Doppler ultrasound for discriminating benign from malignant adnexal masses. We have found that 4D-STIC technology does not improve diagnostic performance of 2D ultrasound. However, we observed that a low vPI increases the probability of malignancy in questionable masses with suspicious morphologic and vascular features in gray-scale and color Doppler ultrasound. We also found that this technique may be time consuming.

The main strengths of our study are three. First, to the best of our knowledge, our study is the first reported one assessing the potential role of 4D-STIC in the differential diagnosis of adnexal masses.

Second, the study is prospective and its design applies a logical sequential approach. Most adnexal masses can be accurately and correctly classified using conventional 2D ultrasound. Therefore, advanced techniques, such as 4D-STIC ultrasound are not needed for these masses. The problem arises for those questionable masses that are difficult to classify using conventional 2D ultrasound. The problem arises for those questionable masses that are difficult to classify using conventional 2D ultrasound. The problem arises for those questionable masses that are difficult to classify using conventional 2D ultrasound. The problem arises for those questionable masses that are difficult to classify using conventional 2D ultrasound.

On the contrary, not all masses included were surgically removed. This means that the most accurate reference standard is lacking in a certain number of cases. This might be a source of bias, since we are assuming that persisting benign appearing masses are benign, which might be not true. However, this limitation could be a strength since, in current practice, conservative management of apparently benign masses is gaining acceptance. Therefore, our data could be interpreted within a more “real-life” clinical setting.

We did not analyze interobserver variability in this study, but previous data suggest that calculation of vPI is reproducible between different observers when they used the same STIC sequence. However, reproducibility of STIC technology when different observers acquire STIC sequence from the same adnexal mass has not been assessed. Additionally, no previous study has been reported. So, we cannot compare our results with any other study.

In summary, 4D-STIC technology does not improve the overall diagnostic performance of conventional 2D ultrasound for discriminating benign from malignant adnexal masses. However, the probability of malignancy of questionable masses increases when a low vPI is found. Further studies are needed for elucidating definitively the role of 4D-STIC in the differential diagnosis of adnexal masses.

**REFERENCES**


29. Valentin L, Ameye L, Jurkovic D, Metzger U, Lécuru F, Van Huffel S, Timmerman D. Which extrauterine pelvic masses are difficult to correctly classify as benign or malignant on the basis of ultrasound findings and is there a way of making a correct diagnosis? Ultrasound Obstet Gynecol 2006 Apr;27(4):438-444.


