Can combination of hysterosalpingography and ultrasound replace hysteroscopy in diagnosis of uterine malformations in infertile women?

Mansoureh Vahdat1, Elaheh Sariri2, Maryam Kashanian3, Zahra Najmi4*, Alireza Mobasser5, Mahjabin Marashi6, Behnaz Mohabbatian7, Shideh Ariana6 Yousef Moradi9

Received: 20 May 2015 Accepted: 28 November 2015 Published: 10 April 2016

Abstract

Background: Müllerian anomalies are associated with infertility. Hysteroscopy as the gold standard for evaluating Müllerian anomalies is an invasive, expensive and risky procedure which requires enough experience. Transvaginal sonography (TVS) and hysterosalpingography (HSG) are less invasive procedures, but there is little known about the accuracy of these tests. The aim of this study was to evaluate the accuracy of the combination of TVS and HSG with hysteroscopy as the gold standard.

Methods: Medical records of infertile women who were undertaken all three diagnostic modalities were reviewed to analyze their sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Results: Ninety-nine infertile women were assessed with a mean±SD age of 29.1±6.47 years, mean±SD duration of the marriage of 8.9±10.28 years, and mean±SD duration of infertility of 5.6±4.16 years. The sensitivity, specificity, PPV, and NPV of TVS were 98.55%, 30%, 76.4%, and 90%, respectively. HSG had a sensitivity of 95.6%, specificity of 60%, PPV of 84.62%, and NPV of 85.71%. When both modalities were combined, the sensitivity, specificity, PPV, and NPV were 94.2, 66.67, 86.67, and 83.33%, respectively. The diagnostic accuracy of single TVS, HSG or combined techniques was statistically similar that was equal to 77.7, 84.8 and 85.8% respectively.

Conclusion: The accuracy of combination of two diagnostic modalities, 2D TVS and HSG is not higher than HSG alone for assessing uterine malformation in infertile women.

Keywords: Ultrasound, Hysterosalpingography, Hysteroscopy, Müllerian anomaly, Accuracy, Infertility.


Introduction

Müllerian anomalies are a group of female reproductive system malformations develop during early embryologic life (1). These congenital abnormalities might present with irregular menses, repeated abortions and infertility, and the diagnosis often delays until adulthood. Infertility af-
Combination of hysterosalpingography and ultrasound for diagnosis of uterine malformations

ffects 9% of women of reproductive age (2). It is estimated that 3.5% of women with infertility have a congenital Mullerian uterine anomaly, a 21-fold increased prevalence compared to normal fertile women (1).

For this high prevalence, one of the first steps in the workup of female infertility is a radiologic evaluation for Mullerian anomalies. The optimal diagnostic test should be accurate, cost-effective, and minimally invasive. Hysterosalpingography (HSG) demonstrates fallopian tubal patency, submucosal fibroid, endometrial polyp, hypertension, hyperkinesia, and intrauterine adhesions and septa. Abdominal ultrasound displays external anatomy of the uterus and fallopian tubes, even in patients with an imperforate hymen, artic vagina, and transverse vaginal septum, without the risk of radiation, contrast allergy, uterine perforation, and infection (3). Hysteroscopy is known as the gold standard method for the visualization of the Mullerian structures and simultaneous treatment. Although hysteroscopy is safe and efficient, this invasive procedure is associated with increased risk of uterine cavity perforation, fluid overloads, bleeding and infection (4-6).

HSG and ultrasound are routinely used in the evaluation of infertility, and the accuracy of each modality in comparison of hysteroscopy as the gold standard method have been evaluated. However, the accuracy of their combination is still indeterminate. So the aim of this study was to compare the diagnostic accuracy of combined transvaginal sonography (TVS) and HSG with that of hysteroscopy in the assessment of uterine anomalies.

Methods

After approval by Ethics Committee of Iran University of Medical Sciences, the study was performed on medical records of infertile women who had diagnostic or therapeutic hysteroscopy at the Department of Obstetrics and Gynecology of Rasoul e Akram Hospital, Tehran, Iran, between Januarys and March 2013. Our study population was infertile women with the report of Mullerian anomalies in their TVS or HSG evaluation that random elected. The analysis was performed on patient’s records which had complete data of prior two-dimensional TVS, HSG and hysteroscopy all done in our center.

Data of a full medical history and complete gynecologic physical examination was extracted from clinical visits for all patients. Infertility was defined as the absence of conception in the at least past 12 months, despite unprotected intercourse. The age, duration of infertility, duration of the marriage, history of fertility or abortion, and menstrual characteristics of patients were recorded. Menstrual bleeding was categorized as irregular if each cycle lasted >35 days or the cycle length varied >10 days. Infertility was categorized as primary or secondary based on the presence of any previous conception. Abortion was defined as spontaneous loss of the fetus before 20 weeks of gestation.

All Participants had undergone TVS, HSG and hysteroscopy examination in our center. Hysteroscopy had performed by experienced gynecologists during the proliferative phase of menstrual cycle, with a 4.5-mm rigid hysteroscopy equipped with a 30 degree lens. Normal saline was used as the distension medium, and a cuff-mounted manometer controlled the pressure between 80 and 100 mmHg. All procedures were undertaken under general anesthesia and without prophylactic antibiotic.

TVS had performed at Radiology Department after proper preparation. HSG had performed and interpreted by gynecologists during the proliferative phase of menstrual cycle. HSG initiates with an injection of 10-ml contrast medium with a sterile cannula into the uterine cavity after placement of a vaginal speculum. Then three plain radiographs were taken to illustrate filled view of the uterine cavity, filled view of fallopian tubes, and contrast medium leak into the abdomen.

Data were analyzed using the SPSS version 19. Results are presented as numbers
and percentages. The mean±SD duration of marriage and infertility were calculated separately. The sensitivity, specificity, and positive and negative predictive values of ultrasound and HSG were compared to hysteroscopy as the gold standard. Accuracy was calculated as:

\[
\text{Accuracy} = \frac{(\text{True Positive} + \text{True Negative})}{(\text{True Positive} + \text{False Positive} + \text{True Negative} + \text{False Negative})}
\]

Also in the study used of LR to calculate the probability of disorder while adapting to varying prior probability of the chance of disease from different contexts.

Likelihood ratio calculated with:

Positive likelihood ratio: Sensitivity / 1 - specificity and negative likelihood ratio: 1 - sensitivity / specificity

### Results

Data collected from 99 patients who underwent all three diagnostic modalities. The mean±SD age, mean±SD duration of marriage, and mean±SD duration of infertility were 29.1± 6.47 years, 8.9±10.28 years, and 5.6±4.16 years, respectively. Of the enrolled patients, 32.4% had, at least, one previous conception. Forty-five patients (45.4%) had an irregular menstrual bleeding history. Only 28 women (28.3%) had the history of abortion. Table 1 shows the demographic data of infertile population who underwent hysteroscopy.

The diagnostic accuracy parameters of TVS, HSG, and the combination of both modalities are shown in Table 2. TVS reported 10 (11.3%) patients as abnormal while only 30 (30.3%) of them were confirmed using hysteroscopy (sensitivity=98.55%, specificity=30%, PPV=76.4%, NPP=90%). HSG detected abnormality in 21 (22%) patients while only 78 (78%) were normal with this test (sensitivity=95.6%, specificity=60%, PPV=84.62%, NPP=85.71%). When both modalities were combined, the sensitivity, specificity, PPV, and NPV were 94.2%, 66.67%, 86.67%, and 83.33%, respectively. Positive likelihood ratios for TVS, HSG, and combined techniques was statistically similar that was equal to 77.7, 84.8 and 85.8 % respectively.

### Table 1. Demographic data of the patients who underwent hysteroscopy

<table>
<thead>
<tr>
<th>Number of patients</th>
<th>99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean±SD Age (year)</td>
<td>29.10± 6.47</td>
</tr>
<tr>
<td>Mean±SD duration of Marriage (year)</td>
<td>8.93± 10.28</td>
</tr>
<tr>
<td>Mean±SD duration of Infertility (year)</td>
<td>5.59± 4.16</td>
</tr>
<tr>
<td>Menstrual bleeding Pattern</td>
<td></td>
</tr>
<tr>
<td>Regular, N (%)</td>
<td>54 (54.6)</td>
</tr>
<tr>
<td>Irregular, N (%)</td>
<td>45 (45.4)</td>
</tr>
<tr>
<td>Infertility kind</td>
<td></td>
</tr>
<tr>
<td>Primary, N (%)</td>
<td>55 (55.6)</td>
</tr>
<tr>
<td>Secondary, N (%)</td>
<td>44 (44.4)</td>
</tr>
<tr>
<td>History of abortion, N (%)</td>
<td>28 (28.3)</td>
</tr>
</tbody>
</table>

### Table 2. Diagnostic accuracy parameters of TVS, HSG, and combination of both modalities for uterine malformations (n=99)

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPP (%)</th>
<th>LR+</th>
<th>LR-</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVS</td>
<td>98.55 (92.19 – 99.99)</td>
<td>30 (14.7 – 49.4)</td>
<td>76.4 (66.22 – 84.22)</td>
<td>90.0 (55.5 – 99.7)</td>
<td>1.41 (1.11 – 1.78)</td>
<td>0.05 (0.01 – 0.36)</td>
<td>77.7 (68.3 – 85.5)</td>
</tr>
<tr>
<td>HSG</td>
<td>95.6 (87.8 – 99.09)</td>
<td>60 (40.6 – 77.3)</td>
<td>84.62 (74.67 – 91.79)</td>
<td>85.71 (63.66 – 96.95)</td>
<td>2.39 (1.54 – 3.72)</td>
<td>0.07 (0.02 – 0.23)</td>
<td>84.8 (76.2 – 91.2)</td>
</tr>
<tr>
<td>TVS+ HSG</td>
<td>94.2 (85.8 – 98.4)</td>
<td>66.67 (47.19 – 82.71)</td>
<td>86.67 (76.84 – 93.42)</td>
<td>83.33 (62.62 – 95.26)</td>
<td>2.83 (1.70 – 4.70)</td>
<td>0.09 (0.03 – 0.23)</td>
<td>85.8 (77.4 – 92.0)</td>
</tr>
</tbody>
</table>

Note: Prevalence=48.5%, the numbers in parentheses are the limits of the 95% confidence interval.

TVS: Transvaginal sonography, HSG: Hysterosalpingography, PPV: Positive predictive value, NPP: Negative predictive value

LR+: Positive likelihood ratio, LR-: Negative likelihood ratio.
**Discussion**

We found that when both modalities (TVS and HSG) were combined, the sensitivity, specificity, PPV, and NPV were 94.2, 66.67, 86.67, and 83.33%, respectively. So HSG and the combination of both diagnostic modalities are accurate enough and are not different for assessing uterine malformation in infertile women. The accuracy of both is higher than TVS alone.

Müllerian anomalies are associated with infertility, so an accurate diagnosis of these defects would result in proper management of infertility (7). Although the gold standard for diagnosis is hysteroscopy, there is a need for a noninvasive, easy and accurate method for evaluating infertile patients with suspected malformations. The aim of this study was to compare the accuracy of the combination of two other methods with hysteroscopy in infertile population.

Ultrasound is the preferred diagnostic test of Müllerian anomalies in some centers because of its several advantages. It is an easy, economical, quick, non-invasive method for visualizing the external contour of uterus and myometrium simultaneously. This radiation-free test is widely available and needs no special preparation. However, the quality of its image easily degraded by bowel gas or abdominal fat pad. TVS can reveal more detailed information (3).

In the present study, ultrasound only identified uterine abnormalities in 19 of 48 defects, missing almost 60%. In a similar study, Caliskan et al. (8) reported the sensitivity and specificity of two-dimensional sonography as 30.2 and 78.1% in the follicular phase (42.1 and 81.2% in the luteal phase, respectively). Soares et al. (9) have reported a sensitivity of 44.4%, specificity of 100%, PPV of 100%, and NPV of 91.8% for TVS in the diagnosis of uterine malformations. On the contrary, TVS has demonstrated 100% sensitivity and 80% specificity in detecting Müllerian anomalies in a study by Pillerito et al. (10). Although the operator's experience and the type of transducer might result in different test accuracy, owning to its poor sensitivity in this series, 2D TVS does not seem like a suitable tool for screening uterine anomalies.

In our study HSG had not similar diagnostic sensitivity to TVS. The overall accuracy of HSG for identifying uterine malformations was almost 84.8%. In a study by Acholonu et al. (11) the HSG had the sensitivity of 58.2% and specificity of 25.6% (accuracy, 50.3%) for detecting Müllerian anomalies in infertile population. Sakar et al. (12) compared the HSG with laparoscopy and reported the sensitivity of 63%, specificity of 89.3%, PPV of 92%, NPV of 55%, and total accuracy of 72% in diagnosing peritoneal factors of infertility. Another study of 216 infertile women presented the sensitivity of 80.3%, specificity of 70.1%, PPV of 84.5%, NPV of 64.5%, and accuracy of 76.4% for HSG in detecting intrauterine abnormalities (13). Brown et al. (14) estimated 60% accuracy of HSG for detecting Müllerian anomalies. Reuter et al. (15) reported the accuracy of differential diagnosis of septate and bicornate uterus HSG interpreted by a radiologist, a gynecologic, and combined with ultrasound examination as 55, 62.5 and 90% respectively. Others have reported the accuracy of 22.2% for HSG (10).

There are reports of more accurate diagnostic methods. Bocca et al. (16) analyzed the accuracy, costs, risks and benefits of 3D TVS and HSG among an infertile population with confirmed congenital Müllerian anomalies and reported the 3D TVS a more accurate, cost-effective, and safer method. Other investigators have also recommended 3D TVS for detecting Müllerian anomalies (8,17). In a study by Faivre et al. (18), 3D TVS was compared with hysteroscopy and MRI in differential diagnosis of bicornate and septate uterus and the reported accuracy was 100, 90 and 77% for each diagnostic method respectively; they concluded that the 3D TVS is a mandatory step in the assessment of patients with suspected anomalies.

The main limitation of our study was its retrospective design. Hysteroscopy is not
usually performed in patients with normal ultrasound or HSG findings, so the diagnostic accuracy of these methods could be easily affected. Because both diagnostic tests are interpreted by physicians, the experience of the clinician influences the results.

**Conclusion**

In conclusion, 2D TVS, HSG, and the combination of both diagnostic modalities are not accurate enough for detecting uterine malformations. Further studies are needed to evaluate the accuracy of 3D TVS and its combinations in the diagnosis of different uterine anomalies.

**References**