Diagnostic Accuracy of Transvaginal Ultrasound in the Evaluation of Female Infertility: A Systematic Review and Meta-analysis

Nur Aisyah Kori, Nur Hayati Jasmin*

School of Medical Imaging, Faculty of Health Sciences, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

Corresponding author: nurhayatijasmin@unisza.edu.my

Received: 3rd August 2023       Accepted: 12th December 2023       Published: 24th December 2023

Abstract
Uterine abnormalities can significantly impact a woman’s fertility and overall gynecological well-being, presenting a diverse range of structural variations that may pose challenges or complexities in both conception and pregnancy. Ultrasound has become an essential component of the modern diagnosis of female infertility and the management of assisted conception to investigate endometrial and uterine factors that affect embryo implantation. This study aims to assess the diagnostic accuracy of TVS in evaluating female infertility and how different types of abnormalities in female infertility affect diagnostic accuracy of TVS. Systematic search was conducted following the guidelines of the Preferred Reporting Items for Systematic Reviews (PRISMA) and Meta-Analyses using electronic databases PubMed and ScienceDirect for papers published between 2012 and 2022. The QUADAS-2 tool was employed to assess bias and applicability concerns. In the initial search strategy, a total of 217 articles were discovered. By applying specific criteria for inclusion and exclusion, five studies involving 1306 patients were included in the final review. Forest plot analysis demonstrated moderate heterogeneity for sensitivity and specificity, \(Q = 10.70\) and \(Q = 32.52\) respectively. A random-effects model demonstrated a moderate overall diagnostic accuracy of TVS in evaluating uterine abnormality with a pooled sensitivity of 0.82 (95% CI: 0.73, 0.91), and a specificity of 0.74 (95% CI: 0.65, 0.83). Findings from this study suggested that TVS offers moderate accuracy in evaluating female infertility, although its accuracy may vary depending on the particular condition being evaluated and other contributing factors. TVS are particularly valuable when used with hysteroscopy, as it demonstrates high accuracy in detecting and characterizing abnormalities.

Keywords
Diagnostic accuracy, ultrasound, female infertility, uterine abnormality

Introduction
Infertility is defined as failure to conceive the desired pregnancy after 12 months of unprotected intercourse. For women aged 35 and above, prompt evaluation is recommended if supported by medical history and physical evidence. Treatment should begin within six months of the initial diagnosis1. Infertility
in women can have various causes, including abnormalities with the fallopian tubes, uterus, cervix, or endometrium. Due to the complexity of abnormalities of the uterus, the diagnostic workup and treatment planning for female infertility are heavily dependent on imaging techniques. In most cases, a clinical evaluation is always followed by an imaging evaluation for female infertility. Hysteroscopy is often used for evaluating the uterine cavity, the lumina and fallopian tube patency, which is one of the most common causes of female infertility. Hysteroscopy is also the preferred method for endoscopic visualization of the uterine cavity as it enables the simultaneous detection and treatment of intracavitary anomalies compared to other methods, making it the gold standard. However, hysteroscopy is contraindicated to patients with early pregnancy, pelvic inflammatory disease and severe allergy to iodine contrast agents and have several limitations that include low specificity for intrauterine pathology and highly invasive procedure.

Ultrasound has become an essential component of the modern investigation of female infertility and the management of assisted conception, which can access endometrial and uterine factors that affect embryo implantation. Fertility scans are required for other gynecologic scan indications, as well as knowledge of basic reproductive physiology, the causes of infertility, and how ultrasound can aid in investigating and treating these issues. Transvaginal ultrasound (TVS) provides most of the diagnostic information in a pelvic scan. Modern ultrasound equipment includes advanced Doppler facilities that improve the evaluation of uterine and endometrial blood flow in assessing endometrial receptivity. In infertile couples, ultrasound is now a critical part of the early investigation and treatment of infertile females. This systematic review and meta-analysis aim to review the diagnostic accuracy of transvaginal ultrasound in evaluating female infertility and to determine the types of uterine abnormalities found in TVS assessment of female infertility.

Materials and Methods

Search Strategy
A comprehensive literature search was performed by an independent author on PubMed and ScienceDirect to search for papers published between 2012 to 2022. The study selection process adhered to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and employing the Population, Intervention, Comparison, Outcome, and Study type (PICOS) assisted in the development of suitable keywords to find relevant publications as shown in Table 1. Studies were screened to meet the predefined inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Population</th>
<th>Women who are unable to get pregnant after 12 months of unprotected intercourse.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>Transvaginal ultrasonography (TVS).</td>
</tr>
<tr>
<td>Comparison</td>
<td>Histopathology and/or Hysteroscopy.</td>
</tr>
<tr>
<td>Outcome i.</td>
<td>Sensitivity, specificity, positive predictive value, and negative predictive value of TVS in the evaluation of female infertility and uterine abnormality.</td>
</tr>
<tr>
<td>Outcome ii.</td>
<td>Types of uterine abnormality.</td>
</tr>
<tr>
<td>Study Type i.</td>
<td>Randomised control studies and quasi-experimental studies.</td>
</tr>
<tr>
<td>Study Type ii.</td>
<td>Observational analytical studies.</td>
</tr>
</tbody>
</table>

Data Extraction and Analysis
Data were extracted into Microsoft Excel for analysis. The methodological quality and risk of bias of included studies were assessed by two authors using the Quality Assessment of Diagnostic Accuracy Studies–2 (QUADAS-2) tool. This tool comprises of 12 tailored signaling questions to evaluate each study’s bias risk and applicability concerns across seven domains. By using QUADAS-2, potential biases and concerns related to diagnostic accuracy evaluation were systematically assessed. Microsoft Excel 2019 was used to generate forest plots for meta-analysis and visualization of the sensitivity and specificity of TVS in
the evaluation of female infertility. Heterogeneity among studies was assessed using the Q statistic, while a random effects model was employed to calculate point estimates and corresponding 95% confidence intervals (CIs). These estimates were presented within the forest plots to provide a comprehensive overview of TVS accuracy measures in female infertility assessment.

Results

Study Selection

Two databases were searched, resulting in 217 initial articles. After eliminating duplicates, 131 studies were assessed based on titles and abstracts. Of these, 32 studies were disqualified for not meeting the specified study types, and 67 lacked the necessary information. Case series, reviews, and descriptive articles were excluded. The remaining 32 studies were further evaluated, with six excluded for not using transvaginal ultrasound as the index test. Thirteen studies did not use histopathology or hysteroscopy as the appropriate reference standard, and six lacked diagnostic accuracy values. The final two studies recurrent pregnancy loss, repeated implantation, and infertile males which were included exclusion criteria of this study. Five included studies were thoroughly reviewed, providing sufficient data for evaluating ultrasound’s diagnostic accuracy in detecting female infertility. Figure 1 shows the summary of the study selection process following the PRISMA guidelines.

Risk of Bias and Applicability Concern

Quality assessment of the 5 included studies is presented in Table 2 and Figure 2. All five papers were conducted by blinding the reviewer to diagnostic test findings. This is important to avoid bias. The studies included in the final review generally displayed good quality research, with a slight risk of bias and concerns about patient selection applicability. Another aspect to address is clinical review bias, which could be influenced by patient characteristics or other test results, was not found in any included studies as shown in Table 2. Overall, most studies displayed low risk of bias and applicability concern.

Study Characteristics

The studies reported the age of included patients, ranging up to 65 years in the study by Maiti et al. (2018), with varying sample sizes. Niknejadi et al. (2012) had the largest sample size of 643 participants, while Choudhary et al. (2017) had 50 participants. Inclusion criteria mainly involved female patients with infertility issues undergoing TVS and hysteroscopy evaluation. Niknejadi et al. (2012) included females aged 20-45 with primary or secondary infertility, while Choudhary et al. (2017) included perimenopausal patients (40-55) with uterine bleeding and uterus less than 12 weeks in size. Two studies had no exclusion criteria, while the other three excluded patients with various conditions. Table 3 summarizes the characteristics of the studies under review.

The studies used different technologies and variations as index tests, leading to varying results and efficiency. For example, Vitner et al. (2013) used a Voluson 730 Pro (GE Healthcare) with a 5-9 MHz transvaginal transducer, and Maiti et al. (2018) used two-dimensional transvaginal ultrasonography. Study settings varied, depending on the type and duration of infertility, previous TVS between 2009-2011, patients admitted to the emergency department or clinic, and perimenopausal patients with specific criteria. The reference standard used in all studies was hysteroscopy. The primary outcome indicated that TVS can be a suitable alternative to hysteroscopy, particularly for diagnostic purposes in women with normal uterine cavities.
Figure 1: Summary of Study Selection Process Following PRISMA Guidelines
Table 2: Results of Quality Assessment for Individual Studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Risk Of Bias</th>
<th>Applicability Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient Selection</td>
<td>Index Test</td>
</tr>
<tr>
<td>Babacan et al. (2014)</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Choudhary et al. (2017)</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Maiti et al. (2018)</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Niknejadi et al. (2012)</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>Vitner et al. (2013)</td>
<td>?</td>
<td>☐</td>
</tr>
</tbody>
</table>

Low Risk ☐ High Risk ☐ Unclear Risk

Figure 2: Graphical Summary of Quality Assessment for Risk of Bias and Concerns Regarding Applicability.
Table 3: Summary of studies using TVS in the evaluation of uterine abnormality in the assessment of female infertility.

<table>
<thead>
<tr>
<th>Studies</th>
<th>Patient Characteristics</th>
<th>N</th>
<th>Age (Range)</th>
<th>Uterine Abnormalities</th>
<th>TVS Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babacan et al. (2014)</td>
<td><strong>Inclusion criteria:</strong> Patients admitted to the outpatient clinic with complaints of uterine bleeding, postmenopausal bleeding and lower abdominal pain.</td>
<td>285</td>
<td>24-89</td>
<td>Endometrial polyp (46.7%) Atrophy (9.5%) Endometritis (4.6%) Uterine myoma (4.2%) Cancer (2.1%) Retained products of conception (1.4%) Hyperplasia (1.1%)</td>
<td>PowerVision 6000 SSA-370A ultrasound equipment with a 5.0 to 7.5 MHz transvaginal transducer. In premenopausal patients, normal limits of anteroposterior diameter of the endometrium were defined as 4-8 mm in proliferative phase, 8-14 mm in the secretory phase and 6-10 mm in the periovulatory phase.</td>
</tr>
<tr>
<td>Choudhary et al. (2017)</td>
<td><strong>Inclusion criteria:</strong> perimenopausal age group (40-55), uterine bleeding, a uterus that is less than 12 weeks in size.</td>
<td>50</td>
<td>40-55</td>
<td>Menorrhagia (50%) Frequent bleeding in (16%) Intermenstrual bleeding in (18%) Heavy prolonged bleeding in (4%) Irregular bleeding in (12%)</td>
<td>5 MHz transvaginal probe and various sonographic parameters such as endometrial thickness, uterine pathology, adnexa, and any other pelvic pathology is noted.</td>
</tr>
<tr>
<td>Maiti et al. (2018)</td>
<td><strong>Inclusion criteria:</strong> Patients with AUB, patients aged up to 65 with postmenopausal bleeding.</td>
<td>200</td>
<td>20-65</td>
<td>Uterine bleeding: Menorrhagia (56%) Postmenopausal bleeding (15.5%) Metrorrhagia (10.5%) Polymenorrhagia (10%) Polymenorrhea (8%)</td>
<td>Two-dimensional transvaginal ultrasonography. In some cases, transabdominal sonography if needed was performed with curvilinear probe of frequency 3 MHz. Endometrial polyps were identified as overgrowth of endometrial gland and stroma, irregular enlarged endometrium complex. Endometrial thickness was noted for</td>
</tr>
<tr>
<td>Study</td>
<td>Inclusion criteria</td>
<td>Exclusion criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niknejadi et al. (2012)</td>
<td>20 to 45-year-old females with primary or secondary infertility for more than 1 year.</td>
<td>Patients with no accurate visualization of the endometrium due to improper time of sonography or due to echogenic endometrium due to bleeding.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Inclusion criteria:**
- Polypoid lesions (50%)
- Long uterine septa (8%)
- Short uterine septa (23%)
- Uterine fibroids (7%)
- Adhesions (5%)
- Endometrial hyperplasia (4%)

**Exclusion criteria:**
- A Voluson 730 Pro (GE Healthcare, Solingen, Germany) was used for ultrasound testing using the 5–9 MHz transvaginal transducer.

Vitner et al. (2013)

**Inclusion criteria:**
- Patients who had previously performed TVS between 2009-2011

**Exclusion criteria:**
- Not stated.

**Inclusion criteria:**
- Polyps (% not mentioned)
- Fibroids (% not mentioned)
- Retained products of conception (% not mentioned)
The diagnostic accuracy of infertility tests varied, with sensitivity ranging from 68.5% to 96.8% and specificity from 60.7% to 100%. The study by Babacan et al. (2014) had the highest sensitivity at 93.3% 15, while the study by Maiti et al. (2018) had the highest specificity at 99.2% 7. None of the studies reported specificities of 100%. Sensitivity data from all studies ranged from 53.8% to 100% with a 95% confidence interval (CI). All five studies reported specificity values between 44.8% and 100%, also with 95% CI.

Positive predictive values (PPV) ranged from 71% to 84%, and negative predictive values (NPV) ranged from 71% to 94%.

Forest plot analysis of sensitivity and specificity of TVS in the evaluation of female infertility is presented in Table 4. Data demonstrated a moderate heterogeneity (Q = 10.703) and (Q = 32.518), respectively. Employing a random-effects model, transvaginal ultrasound (TVS) demonstrated a favorable overall accuracy for evaluating female infertility with a pooled sensitivity of 0.83 (95% CI: 0.68, 0.97), and a specificity of 0.74 (95% CI: 0.61, 0.87). The diagnostic accuracy varied depending on the specific abnormality being assessed.

The available data on the sensitivity and specificity of this method for various conditions is limited. Among the studies that provided data, the sensitivity and specificity of transvaginal ultrasound for endometriosis varied widely, with Maiti et al. (2018) reported 57.57% sensitivity and 100% specificity 7, Niknejadi et al. (2012) reported 88.3% sensitivity and 91.2% specificity 6, and Vitner et al. (2013) reported 44.8% sensitivity and 81.8% specificity 9. Babacan et al. (2014) reported 0% sensitivity for endometriosis 15. For fibroids, Babacan et al. (2014) reported 50% sensitivity 15, Maiti et al. (2018) reported 100% sensitivity and specificity 7, while Niknejadi et al. (2012) reported 89.2% sensitivity and 99.6% specificity 6, and Vitner et al. (2013) reported 85.7% sensitivity and 73.9% specificity 9. Babacan et al. (2014) reported 0% sensitivity for cancer and hyperplasia 15. However, Maiti et al. (2018) reported 100% sensitivity and 96.15% specificity for hyperplasia 7, while Niknejadi et al. (2012) reported 56.2% sensitivity and 99.6% specificity 6. For adhesions, Niknejadi et al. (2012) reported 33.3% sensitivity and 99.8% specificity 9. Lastly, for

Table 4: Forest Plots of Diagnostic Accuracy of TVS in The Evaluation of Female Infertility.

<table>
<thead>
<tr>
<th>Studies</th>
<th>N</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity [95%CI]</th>
<th>Specificity [95%CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babacan et al. 2014</td>
<td>28</td>
<td>Not mentioned</td>
<td>Not mentioned</td>
<td>Not mentioned</td>
<td>Not mentioned</td>
<td>93.3 [82.1 - 100]</td>
<td>60.2 [51.2 - 69.2]</td>
</tr>
<tr>
<td>Choudhary et al. 2017</td>
<td>50</td>
<td>18</td>
<td>8</td>
<td>5</td>
<td>19</td>
<td>78.3 [53.8 - 100]</td>
<td>70.4 [47.1 - 93.7]</td>
</tr>
<tr>
<td>Maiti et al. 2018</td>
<td>20</td>
<td>Not mentioned</td>
<td>Not mentioned</td>
<td>Not mentioned</td>
<td>Not mentioned</td>
<td>69.6 [58 - 81.2]</td>
<td>99.2 [85.4 - 100]</td>
</tr>
<tr>
<td>Niknejadi et al. 2012</td>
<td>64</td>
<td>250</td>
<td>0</td>
<td>60</td>
<td>40</td>
<td>79 [72.1 - 85.9]</td>
<td>82 [75 - 89]</td>
</tr>
<tr>
<td>Vitner et al. 2013</td>
<td>12</td>
<td>53</td>
<td>0</td>
<td>19</td>
<td>4</td>
<td>93 [76.3 - 100]</td>
<td>Q = 10.70; d.f. = 4</td>
</tr>
</tbody>
</table>

n=Sample size; TP=True Positive; FP=False Positive; FN=False Negative; TN=True Negative

**Diagnostic Accuracy of TVS**

The diagnostic accuracy of infertility tests varied, with sensitivity ranging from 68.5% to 96.8% and specificity from 60.7% to 100%. The study by Babacan et al. (2014) had the highest sensitivity at 93.3% 15, while the study by Maiti et al. (2018) had the highest specificity at 99.2% 7. None of the studies reported specificities of 100%. Sensitivity data from all studies ranged from 53.8% to 100% with a 95% confidence interval (CI). All five studies reported specificity values between 44.8% and 100%, also with 95% CI. Positive predictive values (PPV) ranged from 71% to 84%, and negative predictive values (NPV) ranged from 71% to 94%. Forest plot analysis of sensitivity and specificity of TVS in the evaluation of female infertility is presented in Table 4. Data demonstrated a moderate heterogeneity (Q = 10.703) and (Q = 32.518), respectively. Employing a random-effects model, transvaginal ultrasound (TVS) demonstrated a favorable overall accuracy for evaluating female infertility with a pooled sensitivity of 0.83 (95% CI: 0.68, 0.97), and a specificity of 0.74 (95% CI: 0.61, 0.87). The diagnostic accuracy varied depending on the specific abnormality being assessed.
congenital anomalies, Niknejadi et al. (2012) reported 71.67% sensitivity and 99.93% specificity. These findings suggest that transvaginal ultrasound can be a valuable tool for identifying uterine abnormalities, but its accuracy may vary depending on the condition being investigated.

The distribution of uterine abnormalities in female infertility was studied across five articles which encompass 1306 individuals. Endometrial polyps accounted for 43.74% of cases, followed by fibroids at 29.62%, and hyperplasia at 14.43%. Uterine anomalies were found in 5.49% of cases, while endometrial atrophy was present in 2.78%. Endometriosis and adhesions were less frequently observed, with a prevalence of 1.6% and 1.34%, respectively. Cancer and adenomyosis were the least common abnormalities, occurring in only 0.62% and 0.4% of cases. These findings underscore the importance of thorough investigation and diagnosis of uterine abnormalities in female infertility cases to improve treatment and chances of successful conception.

Endometrial polyps and fibroids were the most common abnormalities found across all five articles, emphasizing their significance in infertile females. Figure 3 shows the percentage of uterine abnormalities found by TVS in infertile females. Endometrial polyps are localized echogenic lesions in the endometrial cavity, and fibroids are non-cancerous growths in or around the uterus. Both can affect fertility and require careful consideration during diagnosis and management. The distribution of specific uterine abnormalities may vary based on the study population and methodology. Overall, this highlights the need for a comprehensive approach to diagnose uterine abnormalities in cases of female infertility.

![Figure 3: Percentage of Uterine Abnormalities Found by TVS in Infertile females.](https://example.com/figure3.png)
Discussion

Diagnostic Accuracy of TVS

The pool sensitivity across the studies was reported as 82.4%, with a 95% confidence interval (CI) of [73.2%, 91.5%]. This means that, on average, the transvaginal ultrasound had a moderate-to-high ability to correctly identify positive cases. The pool specificity was reported as 74%, with a 95% CI of [65.1%, 82.9%]. This indicates that, on average, the test had a moderate ability to correctly identify negative cases. This suggests that, on average, TVS performs well in terms of correctly diagnosing the condition of infertile females. The American College of Obstetricians and Gynecologists (ACOG) established a clinical guideline on the evaluation and treatment of infertility, highlighting the effectiveness of transvaginal ultrasound in diagnosing specific factors contributing to female infertility, including uterine abnormalities, tubal pathology, and ovarian disorders. The guidelines also reported that transvaginal ultrasound demonstrates moderate capability in detecting the prevalent causes of infertility in women.

The reported sensitivity and specificity of transvaginal ultrasound (TVS) for assessing female infertility across these studies varied, with sensitivity ranging from 53.8% to 100% and specificity ranging from 44.8% to 100%. The confidence intervals are considerably wide due to the variability in the reported sensitivity and specificity values across different studies. The analysis of data from five studies reveals a notable heterogeneity in the results concerning sensitivity and specificity, except for the specificity study conducted by Vitner et al. in 2013. The accuracy of TVS varied depending on the types of uterine abnormalities causing the infertility. Across the five studies, TVS showed high detection rates for endometrial polyps and uterine fibroids. Sensitivity for endometrial polyps ranged from 44.8% to 88.3%, and specificity from 81.8% to 100%. This contrasts with a study where they reported a high sensitivity of 96% and specificity of 92% in detecting endometrial polyps. For uterine fibroids, review of the five included articles showed a sensitivity ranged from 50% to 100%, and specificity from 73.9% to 100%. This is similar to a study by Campbell et al. (2019) that reported sensitivity ranging from 63% to 97%, and specificity ranging from 76% to 98% for TVS in detecting uterine fibroids.

TVS is a valuable diagnostic tool for identifying these abnormalities, providing a detailed examination of the uterus with high accuracy. High sensitivity means TVS accurately identifies infertility issues, ensuring no crucial factors are missed. Conversely, high specificity ensures TVS accurately rules out infertility in patients, preventing unnecessary treatments. Findings from this review reveal that TVS has high sensitivity and moderate specificity indicating that clinicians can confidently make informed decisions and initiate targeted treatments based on accurate assessments of infertility. Findings from TVS facilitate clear communication with patients, offering precise information about fertility issues and guiding them through treatments with confidence. However, it is crucial to consider factors like the specific conditions and abnormalities being assessed, operator’s expertise, equipment quality, and TVS techniques and parameters when evaluating the diagnostic accuracy of transvaginal ultrasound. For example, the diagnostic accuracy of TVS in detecting uterine cancer and adenomyosis is low, necessitating additional tests for a definitive diagnosis. Greatly enhances its practical usefulness, aiding clinicians in making accurate diagnoses and improving outcomes for individuals facing infertility challenges.

Distribution of Abnormalities in Infertile Females

Endometrial polyps and fibroids are common in infertile females. Endometrial polyps can disrupt the endometrium’s normal development, hormonal balance, and interfere with embryo implantation or sperm transport. Uterine fibroids, non-cancerous growths from the uterus’s muscular wall, can impact fertility by distorting the uterine cavity or obstructing the fallopian tubes, affecting embryo implantation or the endometrial environment. Uterine cancer and adenomyosis are the least common abnormalities in female
infertility. Uterine cancer affects the uterus lining, and fertility preservation techniques may be considered before cancer treatment. Adenomyosis, where tissue grows into the uterus wall, disrupts the uterus’s structure and function, making it less receptive to implantation\(^1\). While TVS can provide valuable information for detecting these conditions, additional tests like biopsies or MRI may be needed for definitive diagnoses, as TVS accuracy is relatively low.

**Quality Assessment**

The evaluation of the five studies on female infertility using TVS indicates overall high quality. However, there are some limitations that could affect reliability. QUADAS-2 was used to assess quality and bias risk in diagnostic accuracy studies. Systematic reviews and meta-analyses are considered high quality due to their rigorous approach, including thorough search strategies, standardized bias assessment, and statistical techniques for data synthesis. Observational studies generally have favorable attributes like large sample sizes and clear inclusion criteria, indicating study quality. However, some studies have limitations like small sample sizes and lack of blinding among interpreters, impacting reliability. Variations in inclusion criteria and imaging protocols may also affect comparability. A notable limitation is a high risk of bias in patient selection, arising from non-randomized or non-representative sampling, inadequate blinding, or subjective judgment. Caution should be taken when interpreting findings due to potential bias distortion. Future studies should aim for rigorous and unbiased patient selection procedures using techniques like randomization, blinding, and transparent reporting. Despite limitations, the quality assessment supports including these studies in a systematic review of TVS’s diagnostic accuracy in female infertility.

**Study Limitations and Recommendations**

One limitation of the study is using only two databases for data collection, which may result in incomplete coverage of relevant literature and potentially overlook important studies in other databases, limiting the comprehensiveness of findings. To address this, expanding the number of databases used is crucial to ensure more comprehensive coverage. Incorporating additional relevant databases, specialized repositories, and platforms will maximize the retrieval of relevant studies. Documenting the search process with specific search terms used promotes transparency and facilitates reporting of the study's methodology. Employing a more extensive range of databases enhances the evidence base, making the results more accurate and reliable.

**Conclusion**

The study highlights that transvaginal ultrasound (TVS) exhibits good diagnostic accuracy in detecting female infertility demonstrating a high sensitivity and moderate specificity. TVS performs well in identifying common abnormalities such as endometrial polyps and uterine fibroids, with high sensitivity and specificity. However, its accuracy in detecting uterine cancer and adenomyosis is low, necessitating additional tests for a definitive diagnosis. Overall, TVS proves to be a valuable tool in diagnosing certain causes of female infertility, but its accuracy can vary depending on the specific condition being assessed and other factors.

**Funding**

The author(s) received no specific funding for this work.

**Acknowledgments**

The authors gratefully acknowledge Universiti Sultan Zainal Abidin (UniSZA) for giving the opportunities and support in conducting this research.
Conflict of Interest Disclosure
No conflict of interest.

References