Diagnostic accuracy of saline infusion sonography in the evaluation of uterine cavity abnormalities prior to assisted reproductive techniques: a systematic review and meta-analyses

S. Seshadri1*, T. El-Toukhy1, A. Douiri2, K. Jayaprakasan3, and Y. Khalaf1

1Assisted Conception Unit, Guys Hospital, 11th Floor, Tower Wing, Great Maze Pond, London SE1 9RT, UK 2Department Of Public Health Sciences, King’s College London, Capital House, 5th Floor, Room 505, 42 Weston Street, London SE1 3QD, UK 3Royal Derby Hospital, Derby & NURTURE, University of Nottingham, Nottingham NG7 2UH, UK

*Correspondence address. E-mail: v9dy@hotmail.com
Submitted on July 14, 2013; resubmitted on August 17, 2014; accepted on October 7, 2014

TABLE OF CONTENTS

- Introduction
- Methods
  - Search strategy
  - Study selection
  - Quality assessment of the studies
  - Data extraction
  - Statistical analysis
- Outcome measures
- Results
- Discussion

BACKGROUND: The diagnostic accuracy of a 2-D transvaginal scan, which is commonly employed to evaluate the regularity and shape of the uterine cavity in subfertile women, is relatively poor compared with other diagnostic modalities like saline infusion sonography (SIS) or hysteroscopy. SIS is a minimally invasive, cost-effective and acceptable diagnostic modality. Therefore the aim of this systematic review was to assess the diagnostic accuracy of SIS in the evaluation of the uterine cavity in subfertile women.

METHODS: A systematic review was conducted of diagnostic studies that compared SIS with hysteroscopy. Twenty relevant studies (including 1645 procedures) were identified and a subsequent meta-analysis was performed. Electronic databases were searched for relevant studies and references of relevant studies were cross checked. Validity was assessed and data were extracted independently by two authors. Heterogeneity was examined, studies were plotted in an ROC area and data were pooled. The main outcome measure was the diagnostic accuracy of SIS in the evaluation of the uterine cavity in subfertile women.

RESULTS: The pooled sensitivity of SIS in the detection of all intrauterine abnormalities was 0.88 (95% confidence interval (CI): 0.85—0.90). The pooled specificity was 0.94 (95% CI 0.93—0.96). The positive and negative likelihood ratios were 20.93 (95% CI: 9.06—48.34) and 0.15 (95% CI: 0.10—0.22), respectively. SIS had good accuracy in the detection of all intrauterine abnormalities (area under the summary receiver operating curve (sROC) = 0.97 ± 0.01). SIS also had a high pooled sensitivity and specificity in the detection of congenital uterine anomalies, 0.85 (95% CI: 0.79—0.90) and 1.00 (95% CI 0.99—1.00), respectively. However the limitations of the review include the heterogeneity amongst the included studies.

© The Author 2014. Published by Oxford University Press on behalf of the European Society of Human Reproduction and Embryology. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
CONCLUSIONS: SIS is a highly sensitive investigative modality and comparable to the gold standard tool, hysteroscopy in the detection of intrauterine abnormalities in subfertile women. SIS is a highly sensitive and specific test in the diagnosis of uterine polyps, submucous myomas, uterine anomalies and intrauterine adhesions and can be used as a screening tool for subfertile patients prior to IVF treatment.

Key words: saline infusion sonography / sonohysterography / hysterosonography / hysteroscopy

Introduction

One in seven couples have difficulty in conceiving and a majority of them undergo in vitro fertilization (IVF) treatment in order to maximize their chances of becoming pregnant (NICE, 2004). Despite significant advances in assisted reproduction technology, the live birth rate after an IVF cycle is around 30% (Andersen et al., 2008). The most critical step for successful IVF outcome is embryo implantation, which is influenced by a positive cross talk between an adequate quality embryo and a receptive endometrium. A favourable uterine cavity environment influences endometrial receptivity and any uterine cavity pathologies in subfertile women can therefore interfere with the implantation process (Cakmak and Taylor, 2011). Uterine cavity abnormalities include polyps, fibroids, adhesions and congenital malformations, all of which could negatively affect the IVF outcome, and could be amenable to surgical treatment albeit based on poor scientific evidence.

Uterine cavity abnormalities are very common particularly in the subfertile population with a prevalence ranging from 11% to up to 45% (Balmaceda and Ciuffardi, 1995; Hinckley and Milki, 2004; Fatemi et al., 2010; Chan et al., 2011). Furthermore, these uterine cavity pathologies could be associated with adverse pregnancy outcomes, such as miscarriage and pre-term delivery, in addition to subfertility (Pritts et al., 2009; Sunkara et al., 2010; Chan et al., 2011). It is therefore imperative that uterine abnormalities are diagnosed accurately as the surgical correction of these abnormalities may potentially improve the prospects of conception and a subsequent successful pregnancy (Mollo et al., 2009; Pritts et al., 2009).

The diagnostic modalities that are commonly employed to evaluate the regularity and shape of the uterine cavity include a conventional 2-D and 3-D transvaginal scan, saline infusion sonography (SIS), hysterosalpingogram (HSG) and hysteroscopy. While the transvaginal scan is generally performed to screen for uterine pathologies, its diagnostic accuracy is poor with pathologies like polyps, adhesions, and submucous fibroids being missed. Recent studies have reported poor sensitivity and positive predictive value (PPV) rates of transvaginal sonography in the detection of polyoid lesions (Ragni et al., 2005; Bingol et al., 2011). SIS is a minimally invasive, cost-effective and acceptable diagnostic modality (Pujar et al., 2010; Hajishaiha et al., 2011). Saline infusion helps to delineate better intracavitary structures, thereby improving the diagnostic accuracy. SIS is virtually devoid of procedural complications (Elsayes et al., 2009) except for potential intracavitary infection, which can generally be avoided if the procedure is performed under aseptic precautions. Hysteroscopy is considered to be the gold standard investigative modality that ascertains the regularity of the uterine cavity by direct visualization of the endometrium. The sensitivity of hysteroscopy is considered to be nearly 100% in the literature (Van Dongen et al., 2007; Almq et al., 2011; Bingol et al., 2011). However, it is an invasive procedure which may require local anaesthetic or sedation in some of the cases and does not allow concurrent assessment of the myometrium and adnexal structures.

Since SIS can assist in the diagnosis of intrauterine pathology, it has been suggested that it should be used as a routine investigative modality to evaluate uterine cavity for all subfertile patients owing to its relatively non-invasive nature and potentially high diagnostic accuracy (Kim et al., 1998; Yauger et al., 2008). While there are many studies reported on the diagnostic ability of SIS in evaluating uterine cavity abnormalities, those studies have reported varying degrees of diagnostic accuracy in relation to different gynaecological conditions, with no systematic reviews summarizing its diagnostic accuracy in the subfertile population. The aim of this systematic review was to evaluate the diagnostic accuracy of SIS in the diagnosis of intrauterine pathologies in subfertile populations in comparison to the gold standard, hysteroscopy.

Figure 1 Search strategy and included studies.

Total number of citations retrieved from electronic searches and examination of reference lists of primary and review articles
Medline-596
Cochrane-22
Embase-940
Cochrane-31
ISI Web of science-111
Total-1700

Citations excluded after screening titles or abstracts n=16512
Full manuscripts retrieved for detailed evaluation n=38
Articles excluded with reasons
- Representative sample population mixed (i.e. premenopausal, history of abnormal bleeding, fertile patients and history of recurrent miscarriages)+9
- Index test different to SIS+3
- Outcome measures different=2
- Abstract and study (n=2) Detailed results not given
- Partial verification bias not avoided (n=2)

Primary articles included in systematic review n=20
<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>No. of women</th>
<th>Type of study</th>
<th>Quality assessment</th>
<th>Completeness of verification</th>
<th>Phase of menstrual cycle when tests performed</th>
<th>Index test</th>
<th>Reference test</th>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alatas</td>
<td>1997</td>
<td>66</td>
<td>Prospective and comparative study</td>
<td>10/11</td>
<td>91%</td>
<td>Not mentioned</td>
<td>TVS, SIS, HSG</td>
<td>Hysteroscopy</td>
<td>Evaluation of sonohysterography in the detection of intrauterine pathology in infertile patients (primary and secondary infertility)</td>
</tr>
<tr>
<td>Ayida</td>
<td>1997</td>
<td>44</td>
<td>Prospective</td>
<td>11/11</td>
<td>100%</td>
<td>Mid cycle or in the early luteal phase</td>
<td>TVS, SIS</td>
<td>Hysteroscopy</td>
<td>Evaluation of diagnostic accuracy of TVS and SIS in infertile women</td>
</tr>
<tr>
<td>Fleischer</td>
<td>1997</td>
<td>100</td>
<td>Prospective consecutive</td>
<td>8/11</td>
<td>73%</td>
<td>Follicular phase (D6–D12)</td>
<td>SIS and tubal assessment</td>
<td>Hysteroscopy and histological findings</td>
<td>Evaluation of diagnostic accuracy of SIS in infertile women</td>
</tr>
<tr>
<td>Darwish</td>
<td>1999</td>
<td>84</td>
<td>Prospective</td>
<td>8/11</td>
<td>73%</td>
<td>Not mentioned</td>
<td>TVS, HSG, SIS</td>
<td>Laparoscopy and hysteroscopy</td>
<td>Evaluation of diagnostic accuracy of HSG and SIS in infertile women</td>
</tr>
<tr>
<td>Gronlund</td>
<td>1999</td>
<td>20</td>
<td>Prospective</td>
<td>9/11</td>
<td>82%</td>
<td>Not mentioned</td>
<td>SIS</td>
<td>Hysteroscopy and surgical biopsy</td>
<td>Evaluation of SIS in both metrorrhagia, recurrent miscarriage and infertility</td>
</tr>
<tr>
<td>Brown</td>
<td>2000</td>
<td>46</td>
<td>RCT</td>
<td>11/11</td>
<td>100%</td>
<td>Follicular phase of the menstrual cycle</td>
<td>HSG, TVS, SIS</td>
<td>Hysteroscopy</td>
<td>Evaluation of diagnostic accuracy of HSG and SIS in infertile women</td>
</tr>
<tr>
<td>Soares</td>
<td>2000</td>
<td>65</td>
<td>Descriptive, prospective study</td>
<td>11/11</td>
<td>100%</td>
<td>First half of the menstrual cycle</td>
<td>HSG, TVS, SIS</td>
<td>Hysteroscopy</td>
<td>Evaluation of diagnostic accuracy of HSG and SIS in infertile women</td>
</tr>
<tr>
<td>Mancini</td>
<td>2002</td>
<td>106</td>
<td>Prospective study</td>
<td>8/11</td>
<td>73%</td>
<td>Early follicular phase</td>
<td>SIS</td>
<td>Hysteroscopy</td>
<td>Evaluation of diagnostic accuracy of SIS in infertile women, abnormal bleeding and post-menopausal bleeding</td>
</tr>
<tr>
<td>Alborzi</td>
<td>2003</td>
<td>186</td>
<td>Prospective</td>
<td>9/11</td>
<td>82%</td>
<td>First half of the menstrual cycle</td>
<td>TVS, SIS and HSG</td>
<td>Hysteroscopy and laparoscopy</td>
<td>Evaluation of diagnostic accuracy of HSG and SIS in infertile women and women with RM</td>
</tr>
<tr>
<td>Guven</td>
<td>2004</td>
<td>93</td>
<td>Prospective and consecutive</td>
<td>5/11</td>
<td>45%</td>
<td>Procedures performed regardless of phase of cycle</td>
<td>TVS, SIS</td>
<td>Hysteroscopy and surgical biopsy</td>
<td>Evaluation of SIS in detection of intrauterine pathology in infertile patients</td>
</tr>
<tr>
<td>Yucebilgin</td>
<td>2004</td>
<td>115</td>
<td>Retrospective</td>
<td>7/11</td>
<td>64%</td>
<td>Not mentioned</td>
<td>SIS</td>
<td>Hysteroscopy</td>
<td>Evaluation of diagnostic accuracy of SIS in infertile women</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Patients</td>
<td>Study Design</td>
<td>Time of Cycle</td>
<td>Imaging Techniques</td>
<td>Procedures</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>----------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radic</td>
<td>2005</td>
<td>68</td>
<td>Prospective</td>
<td>10/11</td>
<td>First half</td>
<td>Hysteroscopy</td>
<td>Evaluation of diagnostic accuracy of SIS in infertile women &amp; RM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ragni</td>
<td>2005</td>
<td>98</td>
<td>Prospective and</td>
<td>8/11</td>
<td>Follicular phase</td>
<td>TVS and SIS</td>
<td>Hysteroscopy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartkowiak</td>
<td>2005</td>
<td>68</td>
<td>Prospective,</td>
<td>11/11</td>
<td>Proliferative phase (7th day)</td>
<td>TVS and SIS</td>
<td>Hysteroscopy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valenzano</td>
<td>2006</td>
<td>54</td>
<td>Retrospective</td>
<td>7/11</td>
<td>Proliferative phase</td>
<td>HSG, TVS and SIS</td>
<td>Hysteroscopy and laparoscopy Evaluation of diagnostic accuracy of TVS and SIS in infertile women</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiran</td>
<td>2008</td>
<td>122</td>
<td>Prospective</td>
<td>8/11</td>
<td>Not mentioned</td>
<td>SIS</td>
<td>Hysteroscopy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Felice</td>
<td>2009</td>
<td>104</td>
<td>Blinded randomized</td>
<td>10/11</td>
<td>Follicular phase (D7–D14)</td>
<td>Group 1- HSG; Group 2- SIS and HSG</td>
<td>Hysteroscopy and laparoscopy Evaluated the diagnostic accuracy of SIS and HSG in the detection of intrauterine and tubal abnormalities in infertile women</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grimbizis</td>
<td>2010</td>
<td>28</td>
<td>Prospective blind</td>
<td>10/11</td>
<td>Not mentioned</td>
<td>SIS</td>
<td>Hysteroscopy and histological diagnosis Evaluated the diagnostic accuracy of SIS in infertile, premenopausal and post-menopausal women</td>
<td></td>
<td></td>
</tr>
<tr>
<td>El-Sherbiny</td>
<td>2011</td>
<td>180</td>
<td>Comparative,</td>
<td>10/11</td>
<td>D5–D10 of menstrual cycle</td>
<td>2-D SIS and 3-D SIS after normal cavity at HSG and TVS</td>
<td>Hysteroscopy Evaluated the diagnostic accuracy of 2-D SIS, 3-D SIS in the detection of intrauterine pathology in infertile patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luciano</td>
<td>2011</td>
<td>58</td>
<td>Retrospective</td>
<td>6/11</td>
<td>Not mentioned</td>
<td>HyCoSy</td>
<td>Hysteroscopy and laparoscopy and histological diagnosis Evaluated the diagnostic accuracy of SIS and HSG in the detection of intrauterine and tubal abnormalities in infertile women</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Methods**

**Search strategy**

We searched the following electronic databases: MEDLINE (1950), EMBASE (1980), CINAHL, Cochrane Central Register of Controlled Trials and Web of Science (1990) from the date of inception to July 2014.


The two subsets were combined with ‘AND’ to generate a subset of citations relevant to our research question. The reference lists of all known primary and review articles were examined to identify cited articles not captured by electronic searches. Articles, which were frequently quoted, were used in the Science Citation Index to identify additional citations. No language restrictions were placed in our searches. Figure 1 outlines the search strategy and the number of included studies.

**Study selection**

Studies were selected if the target population were subfertile women or women undergoing IVF treatment with or without ICSI. The intervention was the evaluation of the diagnostic accuracy of SIS in the detection of

---

**Table II Questions for the QUADAS tool.**

Recommended quality items derived from QUADAS tool (Whiting et al., 2003)

1. Was the spectrum of patients representative of the patients who will receive the test in practice? (representative spectrum)
2. Is the reference standard likely to classify the target condition correctly? (acceptable reference standard)
3. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? (acceptable delay between tests)
4. Did the whole sample or a random selection of the sample, receive verification using the intended reference standard? (partial verification avoided)
5. Did patients receive the same reference standard irrespective of the index test result? (differential verification avoided)
6. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)? (incorporation avoided)
7. Were the reference standard results interpreted without knowledge of the results of the index test? (index test results blinded)
8. Were the index test results interpreted without knowledge of the results of the reference standard? (reference standard results blinded)
9. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? (relevant clinical information)
10. Were uninterpretable/intermediate test results reported?
11. Were withdrawals from the study explained?
intrauterine pathologies compared with the reference (gold) standard which is hysterectomy. Prospective and retrospective comparative studies were included. Randomized, case control, cohort and cross-sectional studies were included. The characteristics of the included studies are outlined in Table I.

Quality assessment of the studies

The quality assessment of the studies was carried out by using the QUADAS tool as suggested by the Cochrane handbook for diagnostic accuracy systematic reviews. The questions outlined in Table II were used for the quality assessment.
assessment. All the manuscripts meeting the selection criteria were assessed for their methodological quality.

Data extraction
Studies were selected in a two-stage process. The titles and abstracts from the electronic searches were initially scrutinized by two authors (S.S., K.J.) and full manuscripts of all citations that were considered likely to meet the predefined selection criteria were obtained. Final inclusion or exclusion decisions were made by at least two authors (S.S., K.J.) on examination of the full manuscripts. In cases of duplicate publications, the most recent or complete publication was used. Disagreement between the two authors was resolved by Y.K. From each study, outcome data were extracted in 2 × 2 tables.

Statistical analysis
The methodological assessment of the included studies was carried out using RevMan 5.1 (Cochrane Collaboration, Oxford, UK). The analysis of the data was done using the Meta-disc version 1.4 (Zamora et al., 2006) to obtain the forest plots and the summary Receiver Operating Characteristic (ROC) curve. We used the fixed effects model (Mantel and Haenszel, 1959) and random effects models as appropriate (DerSimonian and Laird, 1986). Heterogeneity of the exposure effects was evaluated graphically using forest plots (Lewis and Clarke, 2001) and statistically using the $I^2$ statistic to quantify heterogeneity across studies (Higgins and Thompson, 2002). We used the random effects model to calculate the relative risks if the $I^2$ statistic was > 50. Exploration of the causes of heterogeneity was planned using variation in features of population, exposure and study quality. We performed sensitivity analyses where possible and appropriate to address the clinical and methodological variations.

Outcome measures
The primary outcome of interest was the sensitivity, specificity, positive and negative predictive value (PPV and NPV) and pre- and post-test probabilities of the index test (SIS) in the detection of intrauterine abnormalities in infertile patients compared with the reference standard (hysteroscopy). A subgroup analysis was performed to determine the sensitivity, specificity, PPV and NPV of SIS in the detection of intrauterine polyps, submucous myomas, congenital uterine anomalies and intrauterine adhesions. In the subgroup analysis, the index test was reported as being positive when any intrauterine pathology (i.e. intrauterine polyp, submucous myomas, intrauterine adhesions, uterine anomalies) was detected. Secondary outcome measures of interest were the false positive and false negative rates. A study was considered good quality if it scored well on the QUADAS assessment tool.

Results
The process of literature search and identification and selection of studies is summarized in Fig. 1. From a total of 1700 citations, 20 studies were included in the final analysis. The following three studies were excluded as the index test was different to SIS (Shalev et al., 2000: transvaginal sonography compared with diagnostic hysteroscopy; Dzotsenidze et al., 2006: diagnostic accuracy of HSG compared with hysteroscopy; Caliskan et al., 2010: diagnostic accuracy of real-time 3D sonography in the evaluation of congenital mullerian anomalies). Nine studies were included as the representative sample population was mixed (i.e. premenopausal, history of abnormal bleeding, fertile patients and history of recurrent miscarriages) (Romano et al., 1994; Leone et al., 2003; Kelekci et al., 2005; Milingos et al., 2005; Bulletti et al., 2008; Bingol et al., 2011; La Sala et al., 2011; Ludwig et al., 2011; Soguktas et al., 2012). Two studies (Farina et al., 1997; Gera et al., 2008) were excluded as detailed results were not given. Two studies were excluded as the outcome measures were different (Alatas et al., 1998; Yauger et al., 2008). Two studies were excluded as partial verification was not avoided (Kim et al., 1998; Lindheim and Morales, 2002).

The clinical characteristics of the included studies are detailed in Table I. There were fifteen prospective studies (Alatas et al., 1997; Ayida et al., 1997; Fleischer et al., 1997; Darwish and Yussef, 1999; Gronlund et al., 1999; Soares et al., 2000; Alborzi et al., 2003; Mancini et al., 2002; Guven et al., 2004; Radic et al., 2005; Ragni et al., 2005; Bartkowiak et al., 2006; Kiran et al., 2008; Grimbizis et al., 2010; El-Sherbiny and Nasr 2011) and three retrospective studies (Yucebilgin et al., 2004; Valenzano et al., 2006; Luciano et al., 2011) as well as two randomized studies included in the review (Brown et al., 2000; De Felice et al., 2009). The methodological summary of all the included studies has been outlined in Fig. 2. Each methodological item, expressed as a percentage across all the included studies, is outlined in Fig. 3.

The pooled sensitivity of SIS in the detection of all intrauterine abnormalities was 0.88 with a 95% confidence interval (CI) of 0.85–0.90 (Fig. 4). The pooled specificity of SIS in the detection of all intrauterine abnormalities was 0.94 with a 95% CI of 0.93–0.96 (Fig. 4). The positive and negative likelihood ratios were 20.93 (95% CI 9.06–48.34) and 0.15 (95% CI 0.10–0.22), respectively. The likelihood ratio of >10 indicates that the SIS result has a large effect on increasing the probability of disease presence. Figure 5 shows the summary ROC for SIS in the detection of intrauterine abnormalities. The area under the curve was 0.97, indicating SIS as a near perfect test in discriminating normal uterus from those with intracavitary pathology.

The pooled sensitivity and specificity of SIS in the detection of intrauterine polyps were 0.82 (95% CI 0.76–0.86) and 0.96 (95% CI 0.95–0.98), respectively (Fig. 6). The positive and negative likelihood ratios
were 34.66 (95% CI 8.12–147.92) and 0.22 (95% CI 0.13–0.39). The post-test probability for the detection of intrauterine polyps was 0.87 (95% CI 0.79–0.94). The pooled sensitivity and specificity of SIS in the detection of submucous myomas were 0.82 (95% CI 0.69–0.92) and 0.99 (95% CI 0.98–1.00), respectively (Fig. 7). The likelihood ratios were 44.14 (95% CI 17.77–109.64) and 0.26 (95% CI 0.15–0.45). The post-test probability for the detection of submucous myomas was 0.95 (95% CI 0.86 – 1.00). The pooled sensitivity and specificity of SIS in the detection of congenital uterine anomalies were 0.85 (95% CI 0.79–0.90) and 1.00 (95% CI 0.99–1.00), respectively (Fig. 8). The positive and negative likelihood ratios were 53.87 (95% CI 26.78–108.38) and 0.19 (95% CI 0.10–0.35). The post-test probability in the detection of congenital uterine anomalies was 0.99 (95% CI 0.98–1.00). The pooled sensitivity and specificity of SIS in the detection of intrauterine adhesions were 0.82 (95% CI 0.65–0.93) and 0.99 (95% CI 0.98–1.00), respectively (Fig. 9). The positive and negative likelihood ratios were 34.58 (95% CI 16.68–71.70) and 0.36 (95% CI 0.22–0.58). The post-test probability in the detection of intrauterine adhesions was 0.87 (95% CI 0.75–0.99).

Discussion

This is the first systematic review to report on the diagnostic accuracy of SIS in diagnosing intrauterine pathologies which might be associated with adverse reproductive outcomes. Twenty studies were included in this review. The results of the meta-analyses showed that SIS has a high degree of diagnostic accuracy in the detection of all types of intrauterine abnormalities with a sensitivity and specificity of 88 and 94%,
respectively. The diagnostic accuracy of SIS remained high when analysed separately for individual pathologies like endometrial polyps, submucous myomas, intrauterine adhesions and congenital uterine anomalies. In women suspected of having an endometrial pathology on conventional transvaginal scan, SIS can be considered as an alternative to hysteroscopy especially as the specificity reaches close to 1 in detecting intrauterine adhesions, uterine anomalies, endometrial polyps and submucous myomas.

Our results show that SIS is a highly sensitive investigative modality in the detection of intrauterine abnormalities such as uterine polyps and submucous myomas in subfertile women prior to IVF treatment, in comparison to the gold standard, hysteroscopy. This review is supported by the findings of another review (de Kroon et al., 2003) performed in premenopausal women suffering from abnormal uterine bleeding (study population dissimilar to this review), which showed that SIS is highly accurate (sensitivity and specificity of 95 and 88%, respectively) in the evaluation of the uterine cavity. Most of the studies included in our review utilized conventional 2D ultrasound for performing SIS. The sensitivity of SIS may be improved further to reach 100% in detecting intrauterine lesions when 3D SIS technique is used (Sylvestre et al., 2003), equal to that of HSG (100%) in the detection of intrauterine adhesions (Salle et al., 1999). The PPV and NPV for SIS in the assessment of uterine synechiae has reached 100% (La Sala et al., 2011).

This review has shown that SIS is as sensitive and specific as hysteroscopy in detection of intrauterine abnormalities (polyps, submucous myomas, congenital uterine anomalies and adhesions). Given the high prevalence of intrauterine abnormalities in subfertile women and the high diagnostic accuracy of SIS, the clinical application will be for this technique to be used as a first line screening tool in the assessment of subfertile women and reduce the need for invasive diagnostic procedures such as hysteroscopy. Furthermore, SIS has the added benefit of identifying extruterine/adnexal pathology which is not possible in hysteroscopy. If SIS identifies an intrauterine lesion which requires surgical intervention then the women could be recommended for a diagnostic hysteroscopy or if it identifies an extruterine/adnexal problem then the women could be recommended for a laparoscopy.

In direct comparison to hysteroscopy and laparoscopy, SIS has slightly higher false positive rates. However, this is offset by the numerous other advantages of SIS. Although not part of this review, other independent studies have shown SIS to be better tolerated compared with office hysteroscopy ($P < 0.05$) (van Dongen et al., 2008), with a median pain score of 1.6 compared with 3.2 respectively out of 10 (where 0 = no pain and 10 = unbearable pain) ($P < 0.001$) (Rogerson et al., 2002). This is despite the introduction of smaller diameter hysteroscopes and a vaginoscopic approach which eliminates the need for speculum. Furthermore, SIS has a short learning curve and it has also been shown...
to be anywhere between two and nine times cheaper when compared with hysteroscopy (de Kroon et al., 2003). Despite these advantages, SIS has not become standard practice in the assessment of subfertile women possibly due to the fact that it is not readily available in all IVF units. The other reason could be due to the relatively lower diagnostic accuracy from studies using SIS in post-menopausal women where the sensitivity and specificity is not as robust as in premenopausal women. Furthermore, only hysteroscopy allows for concurrent treatment when an intrauterine pathology is identified and women may prefer hysteroscopy compared with SIS due to the ability to undergo diagnosis and treatment in one visit (van Dongen et al., 2011). Evidence also exists that hysteroscopy might have a positive effect on pregnancy rates and therefore may be preferable to SIS as a screening tool in an infertile population (Pundir et al., 2014). Finally, experts may be awaiting better evidence-based data following the advent of newer techniques such as 3D SIS and gel infusion sonography.

However, this review recommends that SIS should become the investigative modality of choice for subfertile women attending an IVF unit as it is of high diagnostic accuracy, better tolerated, easier and quicker to perform, cost effective and decreases the need for invasive investigations such as hysteroscopy. SIS could be an invaluable tool particularly in the developing countries with limited resources.

There are a number of strengths of this review. The studies were homogenous in terms of study population. In addition to overall sensitivity and specificity to intrauterine abnormalities in general, subgroup analysis was also carried out for specific intrauterine conditions such as polyps, submucous fibroids, congenital uterine anomalies and intrauterine adhesions, which are commonly seen in subfertile women. Results of such analysis are helpful both in terms of clinical application and also to provide evidence-based patient information.

However, there were a few limitations to this review. The different study designs contributed to the heterogeneity. The best study design for the evaluation of diagnostic accuracy of any test will be either a cross sectional or a randomized study. Of the 20 studies, 15 studies...
were prospective in design (Alatas et al., 1997; Ayida et al., 1997; Fleischer et al., 1997; Darwish and Yussef, 1999; Gronlund et al., 1999; Soares et al., 2000; Alborzì et al., 2003; Mancini et al., 2002; Guven et al., 2004; Radić et al., 2005; Ragni et al., 2005; Bartkowiak et al., 2006; Kiran et al., 2008; Grimbizis et al., 2010; El-Sherbiny and Nasr, 2011) while three studies were retrospective in design (Yucebilgin et al., 2004; Valenzano et al., 2006; Luciano et al., 2011). Two studies were randomized studies (between SIS and hysteroscopy) (Brown et al., 2000; De Felice et al., 2009). The other reason for the heterogeneity amongst the studies is the timing of the test done in relation to the phase of the menstrual cycle. Eleven studies performed the SIS in the follicular/proliferative phase of the menstrual cycle. Four studies did not mention which phase of the menstrual cycle the index test was performed in. One study mentioned that the SIS was performed regardless of the phase of the menstrual cycle and one study performed the test in the early luteal phase of the cycle. This study shows the high sensitivity and specificity of SIS in the detection of all intrauterine abnormalities in comparison to the gold standard (hysteroscopy). SIS should ideally be performed in the early follicular or proliferative phase of the menstrual cycle, because leaving it to the latter half of the cycle could result in false positives. This could explain some of the disparities in the observed sensitivity and specificity amongst the studies included in this review. On performing a sensitivity analysis with the data obtained from the randomized trials, the sensitivity of SIS in the detection of all intrauterine abnormalities was 91% with a heterogeneity of <50%.

In the future, 3D SIS may replace the conventional transvaginal sonography and 2D SIS in the investigation of subfertile women as studies are beginning to show better sensitivity and specificity. Furthermore, SIS monitored polypectomy without the need for hysteroscopy has also been shown to be feasible (Lee et al., 2006). 3D SIS could well become a confirmatory test rather than just a screening tool and with the additional possibility of either a biopsy or polypectomy in the same sitting, it could well become the gold standard in the near future. However, the current studies are limited in terms of sample size and lack of randomization. Therefore further research in the form of larger randomized controlled trials is needed to evaluate the diagnostic accuracy and technical feasibility of the above mentioned techniques in the detection of intrauterine abnormalities especially uterine anomalies.

This review has achieved its primary objectives by showing that SIS is a highly sensitive and specific investigatory modality in the detection of intrauterine abnormalities in subfertile woman prior to IVF treatment, in comparison to the gold standard, hysteroscopy. In direct comparison to hysteroscopy and laparoscopy, it has slightly higher false positives. However, SIS is better tolerated, more cost-effective, easy and a time efficient investigatory modality that can be carried out to detect intrauterine
pathology in subfertile women prior to IVF treatment (Pujar et al., 2010; Hajishaiha et al., 2011). Further research is needed to evaluate the diagnostic accuracy of other methods like 3D SIS and the feasibility of performing concurrent surgical procedures for the pathologies like polyps that are amenable to treatment.

Acknowledgements
We thank Angela Clough for her valuable guidance and support, and Michael Moxey for his valuable technical support.

Authors’ roles

Funding
None to disclose.

Conflict of interest
None declared.

References