Cochrane Library

Cochrane Database of Systematic Reviews

Three-dimensional saline infusion sonography compared to two-dimensional saline infusion sonography for the diagnosis of focal intracavitary lesions (Review)

Nieuwenhuis LL, Hermans FJR, Bij de Vaate AJM, Leeflang MMG, Brölmann HAM, Hehenkamp WJK, Mol BWJ, Clark TJ, Huirne JAF

Nieuwenhuis LL, Hermans FJR, Bij de Vaate AJM, Leeflang MMG, Brölmann HAM, Hehenkamp WJK, Mol BWJ, Clark TJ, Huirne JAF.
Three-dimensional saline infusion sonography compared to two-dimensional saline infusion sonography for the diagnosis of focal intracavitary lesions.
Cochrane Database of Systematic Reviews 2017, Issue 5. Art. No.: CD011126.
DOI: 10.1002/14651858.CD011126.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

HEADER 1
ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 4
BACKGROUND 5
OBJECTIVES 6
METHODS 6
RESULTS 8
Figure 1. 9
Figure 2. 10
Figure 3. 11
Figure 4. 13
Figure 5. 14
Figure 6. 15
Figure 7. 16
DISCUSSION 17
AUTHORS' CONCLUSIONS 18
ACKNOWLEDGEMENTS 19
REFERENCES 20
CHARACTERISTICS OF STUDIES 24
DATA 44
Test 1. 2D SIS. 44
Test 2. 3D SIS. 44
Test 3. 2D SIS polyps. 44
Test 4. 3D SIS polyps. 44
Test 5. 2D SIS fibroids. 45
Test 6. 3D SIS fibroids. 45
Test 7. 2D+3D SIS polyps. 45
Test 8. 2D+3D SIS fibroids. 45
ADDITIONAL TABLES 45
APPENDICES 46
CONTRIBUTIONS OF AUTHORS 53
DECLARATIONS OF INTEREST 53
SOURCES OF SUPPORT 54
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 54
INDEX TERMS 56

[Diagnostic Test Accuracy Review]

Three-dimensional saline infusion sonography compared to two-dimensional saline infusion sonography for the diagnosis of focal intracavitary lesions

Lotte L Nieuwenhuis[1], Frederik JR Hermans[2], A J Marjolein Bij de Vaate[1], Mariska MG Leeflang[3], Hans AM Brölmann[1], Wouter JK Hehenkamp[1], Ben Willem J Mol[4], T Justin Clark[5], Judith AF Huirne[1]

[1]Department of Obstetrics and Gynaecology, VU University Medical Center, Amsterdam, Netherlands. [2]Department of Obstetrics and Gynaecology, Academic Medical Center, Amsterdam, Netherlands. [3]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands. [4]Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Adelaide, Australia. [5]Birmingham Women's Hospital, Birmingham, UK

Contact address: Lotte L Nieuwenhuis, Department of Obstetrics and Gynaecology, VU University Medical Center, De Boelelaan 1117, Amsterdam, 1081 HZ, Netherlands. ll.nieuwenhuis@vumc.nl.

Editorial group: Cochrane Gynaecology and Fertility Group.
Publication status and date: New, published in Issue 5, 2017.

Citation: Nieuwenhuis LL, Hermans FJR, Bij de Vaate AJM, Leeflang MMG, Brölmann HAM, Hehenkamp WJK, Mol BWJ, Clark TJ, Huirne JAF. Three-dimensional saline infusion sonography compared to two-dimensional saline infusion sonography for the diagnosis of focal intracavitary lesions. *Cochrane Database of Systematic Reviews* 2017, Issue 5. Art. No.: CD011126. DOI: 10.1002/14651858.CD011126.pub2.

ABSTRACT

Background

Focal abnormalities most commonly acquired within the uterine cavity include endometrial polyps (arising from the endometrium) and submucous fibroids (arising from the myometrium). These benign abnormalities can cause several problems, including abnormal uterine bleeding (AUB) and subfertility. Two-dimensional saline infusion sonography (2D SIS) is a minimally invasive test that can be used to diagnose these pathologies, but it is less accurate than hysteroscopy, which is a more invasive procedure by which an endoscope allows direct visualisation of the uterine cavity. Three-dimensional (3D) SIS appears to enhance sonographic visualisation within the uterine cavity, thereby offering a potentially more accurate minimally invasive diagnostic test.

Objectives

Primary objectives

• To evaluate the diagnostic accuracy of 3D SIS (index test 1) compared with 2D SIS for the diagnosis of focally growing lesions (presence or not) in women with AUB or subfertility, with hysteroscopy performed as the reference test**.**

• To evaluate the diagnostic accuracy of 2D+3D SIS (index test 2) compared with 2D SIS for the diagnosis of focally growing lesions (presence or not) in women with AUB or subfertility, with hysteroscopy performed as the reference test**.** In this case, any abnormality on either modality was regarded as a positive result ('OR' approach).

Secondary objectives

• To evaluate the diagnostic accuracy of 3D SIS (index test 1) compared with 2D SIS according to type of abnormality and discrimination between uterine polyps and submucous fibroids in women with AUB or subfertility, with hysteroscopy and histology used as the reference.

• To evaluate the diagnostic accuracy of 2D+3D SIS (index test 2) compared with 2D SIS according to type of abnormality and discrimination between uterine polyps and submucous fibroids in women with AUB or subfertility, with hysteroscopy and histology used as the reference.

Search methods

We searched the following databases: Cochrane Central Register of Studies Online (CENTRAL CRSO), MEDLINE, Embase, PubMed, Cochrane Gynaecology and Fertility Group (CGF) Specialised Register and CGFG Diagnostic Test Accuracy (DTA) Specialised Register, clinicaltrials.gov and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP). Screening reference lists of appropriate studies was also performed. We screened for eligibility all studies identified from inception until March 2016. We performed searches with no date or language restrictions.

Selection criteria

The population of interest consisted of premenopausal women with AUB or subfertility and postmenopausal women with AUB. Diagnostic test accuracy studies, randomised controlled trials (RCTs) and prospective cohort studies were eligible for inclusion if they evaluated the accuracy of both 2D SIS and 3D SIS for the diagnosis of acquired intracavitary abnormalities with hysteroscopy used as the reference standard. In light of the lack of data for 3D SIS, we also included studies that evaluated the accuracy of 3D SIS alone.

Data collection and analysis

Two review authors read all potentially eligible references after performing a first screening by title and abstract (LLN and FJRH). They independently extracted data to construct 2×2 tables from eligible studies and assessed studies for methodological quality using the QUADAS-2 tool (revised tool for quality assessment of diagnostic accuracy studies). To describe and visually present results, we produced in RevMan forest plots showing pairs of sensitivity and specificity together with 95% confidence intervals from each study, as well as raw receiver operating characteristic (ROC) plots. We displayed paired analyses in an ROC plot by linking sensitivity-specificity pairs from each study by using a dashed line. To compare 3D SIS versus 2D SIS, we restricted analyses to studies that provided 2×2 tables for both tests and used the bivariate meta-analysis of sensitivity and specificity.

Main results

Thirteen studies (1053 women) reported the accuracy of 3D SIS for focal uterine abnormalities; 11 of these (846 women) were suitable for meta-analysis, and eight reported accuracy according to the type of focal abnormality. The design of the included studies seems applicable. The main problem involving the quality of included studies is insufficient reporting of study methods, resulting in unclear risk of bias for several of the quality domains assessed. Therefore, we considered the overall quality of the evidence as low. The summary estimate (11 studies reporting absence or presence of abnormality at 3D SIS) for sensitivity was 94.5% (95% confidence interval (CI) 90.6% to 96.9%) and for specificity 99.4% (95% CI 96.2% to 99.9%). Meta-analysis of the eight studies (N = 716) directly comparing 2D SIS versus 3D SIS showed summary sensitivity of 96.9% (95% CI 91.9% to 98.8%) and summary specificity of 99.5% (95% CI 96.1% to 100%) for 3D SIS. For 2D SIS, summary sensitivity was 90.9% (95% CI 81.2% to 95.8%) and summary specificity was 96.3% (95% CI 86.1% to 99.1%). The difference in accuracy between 2D SIS and 3D SIS was non-significant (P values of 0.07 for sensitivity and 0.10 for specificity).

Authors' conclusions

Low-quality evidence suggests that 3D SIS may be very accurate in detecting intracavitary abnormalities. Meta-analysis revealed no statistically significant differences between 2D SIS and 3D SIS. Summary sensitivity and summary specificity are higher for 3D SIS, but margins of improvement are limited because 2D SIS is already very accurate. When the technology and appropriate expertise are available, 3D SIS offers an alternative to 2D SIS. Both 2D SIS and 3D SIS should be considered alternatives to diagnostic hysteroscopy when intracavitary pathology is suspected in subfertile women and in those with abnormal uterine bleeding.

PLAIN LANGUAGE SUMMARY

Is three-dimensional saline infusion sonography (3D SIS) better than two-dimensional (2D) SIS for detecting polyps and fibroids?

Review question

Is three-dimensional saline infusion sonography (3D SIS) better than two-dimensional (2D) SIS for detecting polyps and fibroids?

Background

The womb (uterus) is one of the female reproductive organs. Inside the cavity of the womb, abnormalities such as polyps and fibroids can grow. Polyps and fibroids can cause problems such as abnormal menstrual bleeding and difficulty getting pregnant. The presence of these polyps and fibroids may be a reason for clinicians to start drug therapy or remove the polyps and fibroids during surgery.

Ultrasonography can provide a picture of the womb and of possible fibroids or polyps. Saline or gel inside the cavity of the womb makes the ultrasound image more clear. This technique is called saline infusion sonography (SIS). Usually, this picture is only two-dimensional. Nowadays, it is possible to make a three-dimensional picture so the type of abnormality can be better seen.

Study characteristics

Review authors searched for studies published from inception until March 2016 and found 13 studies (in total 1053 women), eight of which directly compared 3D SIS versus 2D SIS. Data included all women reporting abnormal menstrual bleeding or difficulty getting pregnant. The number of patients in these studies varied from 23 to 180 women.

Quality of the evidence

In all studies, researchers checked the results of 2D SIS and 3D SIS against results obtained when a camera was used to look inside the womb (hysteroscopy); this is expected to give the true picture but is also more painful for the patient. All studies were performed in the usual way. Some studies did not report several items that might have influenced the results. For example, not all studies made it clear that the person evaluating the ultrasound pictures was unaware of the hysteroscopy results, and vice versa. The main problem involving the quality of included studies is insufficient reporting of study methods, resulting in unclear risk of bias for several of the quality domains assessed. Therefore, review authors considered the overall quality of the evidence as low.

Key results

Low-quality evidence suggests that 3D SIS may be very accurate in detecting polyps and fibroids. Our analysis revealed no clear differences between 2D SIS and 3D SIS. Summary results are higher for 3D SIS but margins of improvement are limited because 2D SIS is already very accurate. Results show that 2D SIS missed a fibroid or polyp in 9 of 100 women and 3D SIS missed a polyp or fibroid in 3 of 100 women who had them. In 4 of 100 women, 2D SIS indicated the presence of polyps or fibroids when there were none, and in less than 1 in 100 women, 3D SIS was wrong. In theory, if both tests were used in a group of 1000 women with abnormal menstrual bleeding, 300 with fibroids or polyps, 27 of the 300 women with polyps/fibroids will be missed by 2D SIS, and 9 of 300 will be missed by 3D SIS.

3D SIS is an alternative to 2D SIS for which the technology and appropriate expertise are available. Both 2D SIS and 3D SIS should be considered alternatives to diagnostic hysteroscopy when intracavitary pathology is suspected in subfertile women and in those with abnormal uterine bleeding.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings table

Study question	Accuracy of 3-dimensional (3D) saline infusion sonography (SIS) compared with 2-dimensional (2D) SIS for the diagnosis of focal intracavitary abnormalities
Patient population	Premenopausal women with abnormal uterine bleeding or subfertility and postmenopausal women with abnormal uterine bleeding; 7 studies included patients with abnormal uterine bleeding, 3 included patients with subfertility and 3 included both types of patients
Prior testing	Eight of 13 studies reported a prior test. Prior tests reported were 2D ultrasonography and hysterosalpingography
Index tests	The review includes studies evaluating the diagnostic accuracy of 3D SIS (index test 1) and studies evaluating the diagnostic accuracy of 2D SIS+3D SIS (index test 2) in comparison with 2D SIS (comparator test)
Reference standard	Hysteroscopy was the reference standard
Target condition	Lesions focally growing inside the uterine cavity (anomalies of the uterine cavity were excluded)
Studies	The search included studies from inception until March 2016. Thirteen studies (1053 women) matched the inclusion criteria and were included for qualitative synthesis: 1 randomised controlled trial (RCT) and 12 prospective cohort studies. Eleven studies (846 women) reported accuracy in detecting presence/absence of an abnormality, and 8 studies reported presence/absence of a specific abnormality (uterine polyp or submucous fibroid). Study size ranged from 23 to 180 participants. Prevalence of the target condition ranged from 14% to 96%
Methodological quality concerns	The design of the included studies seems applicable. The main quality problem with the included studies was insufficient reporting of methods, resulting in unclear risk of bias for several of the quality domains assessed. Therefore, review authors considered the overall quality of the evidence as low
Main outcome	The summary estimate (11 studies reporting 3D SIS) for sensitivity was 94.5% (95% confidence interval (CI) 90.6% to 96.9%) and for specificity 99.4% (95% CI 96.2% to 99.9%) evaluated against hysteroscopy Meta-analysis of the 8 studies (N = 716) directly comparing 2D SIS vs 3D SIS showed no statistically significant difference (P values of 0.07 for sensitivity and 0.10 for specificity). Summary sensitivity of 3D SIS was approximately the same as in the complete set of 11 3D SIS studies: sensitivity 96.9% (95% CI 91.9% to 98.8%); specificity 99.5% (95% CI 96.1% to 100%). The summary sensitivity for 2D SIS was 90.9% (95% CI 81.2% to 95.8%) and for specificity 96.3% (95% CI 86.1% to 99.1%) To characterise the usefulness of the test in different prevalence scenarios, we calculated post-test probabilities (PPVs) for 3 different values of prevalence: 15%, 50% and 90%. PPV would be 96.0%, 99.3% and 99.9%, respectively. Sensitivity analyses showed nearly no influence on the summary estimates of sensitivity and specificity
Conclusion	Low-quality evidence showed 3D SIS to be very accurate in detecting intracavitary abnormalities. Meta-analysis showed no statistically significant differences between 2D SIS and 3D SIS. Summary sensitivity and specificity are higher for 3D SIS but margins of improvement are limited because 2D SIS is already very accurate. 3D SIS is an alternative to 2D SIS for which the technology and appropriate expertise are available. Both 2D SIS and 3D SIS should be considered alternatives to diagnostic hysteroscopy when intracavitary pathology is suspected in subfertile women and in those with abnormal uterine bleeding

BACKGROUND

Target condition being diagnosed

The most common focal intracavitary uterine abnormalities are acquired benign formations arising from the endometrium (endometrial polyps) or from the underlying myometrium (submucous fibroids). These abnormalities are important because they are thought to cause abnormal uterine bleeding (AUB) and subfertility (Golan 2001; Klatsky 2008; Munro 2011; Pritts 2009). It is unclear how focal disruption of the endometrial lining and uterine cavity causes AUB and subfertility. In the case of submucous fibroids, AUB is thought to arise through enlargement of the endometrial surface or as the result of bleeding from stretched and fragile blood vessels (Patterson 1994; Stewart 1996).

Abnormal uterine bleeding affects about 10% to 35% of healthy premenopausal women (Gath 1987; Liu 2007; Santer 2005). Apart from hormonal imbalance, intracavitary abnormalities are the leading cause (Emanuel 1995; Tur-Kaspa 2006; Werbrouck 2011). Postmenopausal bleeding (PMB) is less common than AUB but should always be evaluated to rule out endometrial carcinoma or endometrial hyperplasia. Polyps and submucous fibroids are common in both premenopausal and postmenopausal women, although polyps appear to be more prevalent and submucous fibroids less prevalent in women of post-reproductive age than in those of reproductive age. In a prospective study of more than 1000 women with AUB, 18.4% of premenopausal versus 37.7% of postmenopausal women were diagnosed with polyps, and 14.2% versus 6.2% showed intracavitary fibroids, respectively (van den Bosch 2015). These focal lesions are also highly prevalent in women with subfertility. In a Cochrane review, the overall prevalence of intracavitary abnormalities in women before in vitro fertilisation (IVF) or intracytoplasmic sperm insemination (ICSI) was 11% (Fatemi HM 2010).

Endometrial polyps and submucous fibroids can be diagnosed via transvaginal sonography (TVS). Saline infusion sonography (SIS) is a sonographic technique whereby saline is instilled through the cervical canal and into the uterine cavity during TVS to distend the uterine cavity, thereby enhancing visualisation. Gel can be used as an alternative distension fluid to saline (GIS). Hysteroscopy is an endoscopic procedure whereby the uterine cavity is visualised directly; it is considered the gold standard for detecting polyps and submucous fibroids (Tarneja 2002). Diagnoses obtained during hysteroscopy are more accurate than those obtained by hysteroscopy with histology in both premenopausal and postmenopausal women (Metello 2008; van Dongen 2007). However, hysteroscopy is considered a more invasive diagnostic modality than TVS or SIS and may require use of a general anaesthetic.

Index test(s)

During saline infusion sonography, fluid (saline or gel) is instilled transcervically into the uterine cavity to provide enhanced visualisation of the endometrial lining during TVS examination.

Two-dimensional saline infusion sonography (2D SIS) and diagnostic hysteroscopy are techniques used for detection of intrauterine abnormalities (Dijkhuizen 2003). Both SIS and GIS are simple, safe, well tolerated and accurate for assessment of intrauterine abnormalities (Beemsterboer 2008; Bij de Vaate 2010). Although 2D SIS has become the diagnostic test of choice in most clinical practices, diagnostic hysteroscopy (with histology) remains the gold standard to confirm the presence or absence of an intrauterine abnormality.

Three-dimensional (3D) SIS enhances visualisation of the uterine cavity and is reported to be highly accurate (Abou Salem 2010; Lee 2006; Salim 2005; Terry 2009). Three-dimensional SIS allows examination of the uterus from any angle and in any plane, and this allows the examiner to more precisely measure fibroid size and the extent of protrusion of submucous fibroids into the uterine cavity. This procedure can be performed during the same session as 2D SIS, and the procedure is similar, apart from the requirement that a 3D ultrasonic probe be used. The duration of performing 3D SIS is similar to that for 2D SIS. In 3D SIS, a video can be stored and (re-)studied at any time and in any plane at the ultrasonography machine or on a personal computer. As 3D SIS is reported to be accurate, it may provide additional value over 2D SIS, or may even replace it.

This review focused on studies in which 3D SIS (index test 1) and 2D SIS+3D SIS (index test 2) were compared with 2D SIS (comparator test), using hysteroscopy as the reference test, to identify acquired intrauterine abnormalities, polyps and submucous fibroids. When possible, 3D SIS was compared with 2D SIS according to type of abnormality (i.e. to differentiate between polyps and submucous fibroids) with hysteroscopy and histology as the reference test.

Clinical pathway

In most clinics in developed countries, women with AUB or subfertility undergo TVS. When an intracavitary abnormality on TVS is suspected, a 2D SIS can be planned to diagnose presence, size and type of abnormality (Guideline ACOG; Guideline NVOG; Guideline SOGC). Although hysteroscopy is a relatively safe operation associated with minor complications, it is more expensive and burdensome for the patient than is SIS (Dijkhuizen 2003; Dongen 2011; van den Bosch 2008; Widrich 1996). The main reason it is more burdensome is that it is reported to be more painful than SIS and sometimes requires the use of general anaesthesia.

If an intracavitary abnormality is seen during 2D SIS, a treatment plan can usually be made that involves scheduling an operative hysteroscopic procedure at the point where the abnormality will be resected or morcellated in an outpatient clinic (most polyps) or in a hospital day care setting with the patient under regional or general anaesthesia (most submucous fibroids).

If a normal cavity is seen at 2D SIS, women with AUB are treated first with expectant management or hormonal therapy. Among postmenopausal women with uterine bleeding, the main focus of testing is to exclude endometrial cancer or precancer (atypical endometrial hyperplasia). Four types of diagnostic tests are used: sonographic measurement of endometrial thickness, endometrial sampling, hysteroscopy and SIS. Consensus regarding the sequence in which these methods should be employed in women with PMB is lacking (van Hanegem 2011). However, most guidelines recommend TVS as the first-line test, and that women with an endometrial thickness greater than 4 mm should undergo endometrial sampling. If histological diagnosis is inconclusive, or if PMB recurs, guidelines recommend that an outpatient

hysteroscopy with concomitant endometrial biopsy should be planned (Guideline ACOG; Guideline NVOG; Guideline RCOG).

Rationale

Polyps and submucous fibroids are associated with AUB and subfertility. SIS is a minimally invasive, cost-efficient, outpatient test that can detect these focally growing lesions with a good degree of accuracy (Bij de Vaate 2010; de Kroon 2003) without the need for more invasive and costly hysteroscopy. Detected lesions can be removed subsequently in the hope of alleviating AUB symptoms, excluding malignant or premalignant disease (van Hanegem 2016) and optimising fertility (Bosteels 2015). A meta-analysis of the accuracy of 2D SIS (with hysteroscopy as the gold standard) in AUB revealed pooled sensitivity and specificity of 95% and 88%, respectively (de Kroon 2003). High sensitivity in 2D SIS means that additional diagnostic methods (such as diagnostic hysteroscopy) may be avoided if a normal uterine cavity is detected with SIS. However, 5% of uterine cavity abnormalities were missed and 12% of women without abnormalities were scheduled for unnecessary hysteroscopy as the result of a false-positive test result (they were found to have a normal uterine cavity).

Saline infusion sonography has been performed conventionally with 2D ultrasonic imaging. However, 2D SIS may not be fully accurate in assessment of the type and classification of intrauterine abnormalities (Kroon 2006). Three-dimensional SIS appears to enhance visualisation of the uterine cavity and may help discriminate between polyps and submucous fibroids. Furthermore, 3D SIS can very accurately measure the size and extent of protrusion of submucous fibroids into the uterine cavity (Lee 2006; Mavrelos 2011); these parameters are important determinants for the planning of hysteroscopic procedures in terms of the need for pharmaceutical endometrial downregulation, appropriate treatment setting, type of anaesthesia, choice of surgical instruments and required surgical expertise (Betjes 2009; Wamsteker 1993). Another benefit of an accurate imaging diagnosis is the potential to skip a surgical procedure and provide medical therapy (e.g. oral contraceptives, levonorgestrel-releasing intrauterine system). In the case of fibroids, medical therapy such as a gonadotropin-releasing hormone agonist or a selective progesterone receptor modulator can be considered before surgery (Lethaby 2002; Sancho 2016). Enhanced detection and classification of focally growing intracavitary lesions might result in improved patient outcomes as a result of better treatment and reduced morbidity from unnecessary hysteroscopies or missing abnormalities. Despite the high accuracy reported for 3D SIS in the diagnosis of uterine abnormalities, 2D SIS is usually used in clinical practice because consistent evidence of the diagnostic benefit of 3D imaging and availability of 3D ultrasonic probes are insufficient.

OBJECTIVES

Primary objectives

- To evaluate the diagnostic accuracy of 3D SIS (index test 1) compared with 2D SIS for the diagnosis of focally growing lesions (presence or not) in women with AUB or subfertility, with hysteroscopy performed as the reference test.
- To evaluate the diagnostic accuracy of 2D+3D SIS (index test 2) compared with 2D SIS for the diagnosis of focally growing lesions (presence or not) in women with AUB or subfertility, with hysteroscopy performed as the reference test. In this case, any abnormality on either modality was regarded as a positive result ('OR' approach).

Secondary objectives

- To evaluate the diagnostic accuracy of 3D SIS (index test 1) compared with 2D SIS according to type of abnormality and discrimination between uterine polyps and submucous fibroids in women with AUB or subfertility, with hysteroscopy and histology used as the reference.
- To evaluate the diagnostic accuracy of 2D+3D SIS (index test 2) compared with 2D SIS according to type of abnormality and discrimination between uterine polyps and submucous fibroids in women with AUB or subfertility, with hysteroscopy and histology used as the reference.

METHODS

Criteria for considering studies for this review

Types of studies

All diagnostic test accuracy studies, randomised controlled trials and prospective cohort studies for which a 2×2 contingency table could be reproduced in which 2D SIS and 3D SIS were evaluated with results of hysteroscopy as the reference standard were eligible for inclusion in the review. All studies in which 2D SIS or 3D SIS alone was evaluated were considered eligible for inclusion, although if enough studies (10) were found, we preferred studies that reported both 2D SIS and 3D SIS. We included comparative studies if both 2D SIS and 3D SIS were performed in the same setting, regardless of performance sequence, to avoid a difference in performance between index tests arising from differences among participants and settings. We excluded case control, case report and retrospective cohort studies. We contacted the authors of unpublished studies (only congress abstract or published protocol was available) to facilitate inclusion of additional useful data. We did not apply language restrictions.

Participants

Populations of interest were premenopausal women with AUB or subfertility and postmenopausal women with AUB.

Index tests

We included studies comparing the diagnostic accuracy of 3D SIS alone (index test 1) or 2D+3D SIS (index test 2) versus 2D SIS (comparator test). We defined SIS as positive for a focal intrauterine lesion when any distortion of the endometrial lining was visualised (see below). We preferred that time between index test and reference standard was less than one month.

Target conditions

The target condition was the presence (or absence) of a focally growing lesion in the uterine cavity (endometrial polyp, submucous fibroid). We differentiated between polyps and fibroids by classifying smoothly margined echogenic masses with a homogenous texture as polyps (Parsons 1993), and well-defined round lesions within the myometrium or attached to it, often showing shadows at the edge of the lesion and/or internal fan-shaped shadowing, as uterine fibroids (van den Bosch 2015a). Typically, echogenicity varies and some hyperechogenicity may be present internally. We classified submucous fibroids into types

of fibroid (types 0 to 2) using the International Federation of Gynecology and Obstetrics (FIGO) PALM-COEIN classification for abnormal uterine bleeding: type 0 = pedunculated intracavitary, 1 = submucosal < 50% intramural, 2 = submucosal ≥ 50% intramural (Munro 2011). We did not consider hyperplasia, adhesions and congenital anomalies as focally growing lesions.

Reference standards

Diagnostic hysteroscopy was the reference standard for confirmation of the absence or presence of the target condition. Absence of an intracavitary abnormality was seen as clear vision of the entire cavity without disruption of the endometrial lining. When reported, we also used histology as a reference standard to differentiate the type of abnormality (polyp or fibroid). Because the sensitivity for 2D SIS is known to be high (de Kroon 2003), it might be considered unethical to perform hysteroscopy in women with a negative SIS test (meaning no suspicion of abnormality). To optimise data quantity, we included data from studies with partial verification bias (i.e. when some women did not undergo a hysteroscopy (reference standard) after a negative test result with 2D SIS).

Search methods for identification of studies

Electronic searches

We performed searches in consultation with the Cochrane Gynaecology and Fertility Group (CGFG) Information Specialist without language restrictions from inception until 1 March 2016.

We searched the following electronic databases.

- On the web: Cochrane Central Register of Controlled Trials (CENTRAL) via the Cochrane Register of Studies Online (CRSO) (Appendix 1). We searched the trial registries of clinicaltrials.gov (Appendix 2) and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (Appendix 3). We searched PubMed (Appendix 4) to find published trials not yet indexed in MEDLINE.
- Ovid: MEDLINE (Appendix 5) and Embase (Appendix 6) databases.
- ProCite, the CGFG Specialised Register for Randomised Controlled Trials (RCTs) (Appendix 7) and the CGFG Diagnostic Test Accuracy (DTA) Specialised Register (Appendix 8).

Searching other resources

We carried out a handsearch by screening the reference lists of all included articles. We screened available online conference abstracts in the field of gynaecology.

Data collection and analysis

Selection of studies

Two review authors (LLN and FJRH) independently read all potential studies after first screening by title and abstract (also performed independently by LLN and FJRH). We resolved disagreements on decisions for inclusion through discussion; when agreement was not reached, we consulted a third review author (HAMB). We considered all prospective studies comparing 2D SIS and 3D SIS with hysteroscopy as eligible. We used ENDNOTE as our bibliographic management system. We removed duplicates after checking each study by hand and verified all studies in multiple papers of the same study.

Data extraction and management

Two review authors (LLN and FJRH) independently extracted data from eligible studies - all written in English. We resolved disagreements by discussion between review authors; when agreement was not reached, we consulted a third review author. We contacted study investigators to request additional data on methods and results when we noted missing information or inconsistencies.

Assessment of methodological quality

Two review authors (LLN and FJRH) worked independently to assess studies for methodological quality using QUADAS-2 (revised tool for quality assessment of diagnostic accuracy studies) (Whiting 2011). The QUADAS-2 tool consists of four different domains: patient selection, index test, reference standard, flow and timing. Each domain comprises questions used to assess risk of bias and to address applicability concerns. We supplemented the QUADAS-2 tool with review-specific questions (Table 1). For example, in the index test domain, we added the question whether the level of experience of the (index) test performer was reported. We added to the reference test and target condition domain the question whether the target condition was specified and in which categories. To assess quality, we scored all four domains as having low, unclear or high risk of bias and low or high concern regarding applicability.

Statistical analysis and data synthesis

Data synthesis

The main objective of this review and meta-analysis was to assess the accuracy of (2D+)3D SIS compared with 2D SIS in detecting focally growing lesions. We expected to include studies that evaluated 2D SIS and 3D SIS against the reference standard, as well as studies that evaluated the combination of 2D SIS and 3D SIS against the reference standard. If we found fewer than 10 of these latter studies, we included studies that evaluated 2D SIS or 3D SIS alone.

For studies in which multiple index tests (2D SIS, 2D+3D SIS and 3D SIS) were performed, we constructed a series of 2×2 contingency tables that combined results of investigations provided that they were derived from the total study population and that the definition of a positive result was given for one of the tests.

A secondary objective was to assess the accuracy of index tests according to type of abnormality, while differentiating between polyps and fibroids. In this case, we used the same analyses as described below but the target condition was polyp (vs no polyp) or submucous fibroid (vs no submucous fibroid), depending on what was reported. We wanted to differentiate between polyps and fibroids but could not prepare a 2×2 table because several results were possible. Therefore, in the chosen analyses, we did not directly differentiate between polyps and fibroids but instead analysed them as polyp (vs no polyp) or submucous fibroid (vs no submucous fibroid).

Statistical analysis

To describe and visualise the data, we produced in RevMan forest plots showing pairs of sensitivity and specificity together with

95% confidence intervals from each study, as well as raw receiver operating characteristic (ROC) plots for 2D SIS and for 3D SIS. We displayed paired analyses (studies that tested both 2D SIS and 3D SIS) in an ROC plot by linking sensitivity-specificity pairs from each study with a dashed line (Leeflang 2008).

To compare 3D SIS versus 2D SIS, we restricted our analyses to studies that provided 2×2 tables for both tests, and we used the bivariate meta-analysis of sensitivity and specificity, while allowing for variation in variance of the logit sensitivity and specificity (Reitsma 2005). Even with few studies and/or sparse data (owing to 100% sensitivity or specificity), the advice is to use a hierarchical model (Takwoingi 2015). In cases of fewer than five studies, or when the models did not converge, we planned to use a univariate random-effects logistic regression model.

To analyse the addition of 3D SIS to 2D SIS, we included studies that reported accuracy for the combination of these two techniques. In this case, we regarded as a positive result any abnormality seen on either of the two modalities ('OR' approach). This means that both tests do not need to be positive, if either of the two tests (2D SIS or 3D SIS) was positive and the other negative we considered the test positive. If both tests were negative we considered the test negative.

For 3D SIS, we noted minimal variation in specificity and quite some variation in sensitivity. The models converged only if we assumed no correlation between logit sensitivity and logit specificity, probably because results showed very little heterogeneity, with almost all studies reporting specificity of 100%. If sensitivity varies and specificity does not, the assumption of no correlation may be valid.

We presented results as summary sensitivity and summary specificity. We used the graphical display of a false-positive (1 - specificity) versus a sensitivity plot (ROC plot) showing individual study results, including individual study estimates, the summary operating point (summary values for sensitivity and specificity) and the 95% confidence region on the operating point.

Investigations of heterogeneity

We addressed heterogeneity by adding variables to the bivariate model as covariates if both subgroups included at least three studies. We performed these analyses for clinical symptoms (bleeding vs subfertility); prior testing (prior testing or not); and whether evaluation of 2D or 3D SIS was blinded for clinical information. We also wanted to evaluate the effect of menopausal state (premenopausal or postmenopausal), but the models did not converge when we added this information as a covariate. We also assessed heterogeneity with forest plots and ROC plots.

Sensitivity analyses

We conducted sensitivity analyses to determine whether the methodological characteristics (as assessed by domains of the QUADAS-2 tool) of included studies influenced summary estimates of sensitivity and specificity. We evaluated what would have happened if we would have removed studies at high risk of bias for either of the QUADAS-2 domains. In case of missing or uninterpretable index tests (failure of 2D SIS or 3D SIS), data were classified as positive test results.

Assessment of reporting bias

We aimed to minimise potential impact by ensuring a comprehensive search for eligible studies and by staying alert for duplication of data. We applied no language restrictions and required no translations. We contacted the authors of unpublished studies to ask for data and full text; unfortunately, this information was not available. We recorded ongoing studies.

RESULTS

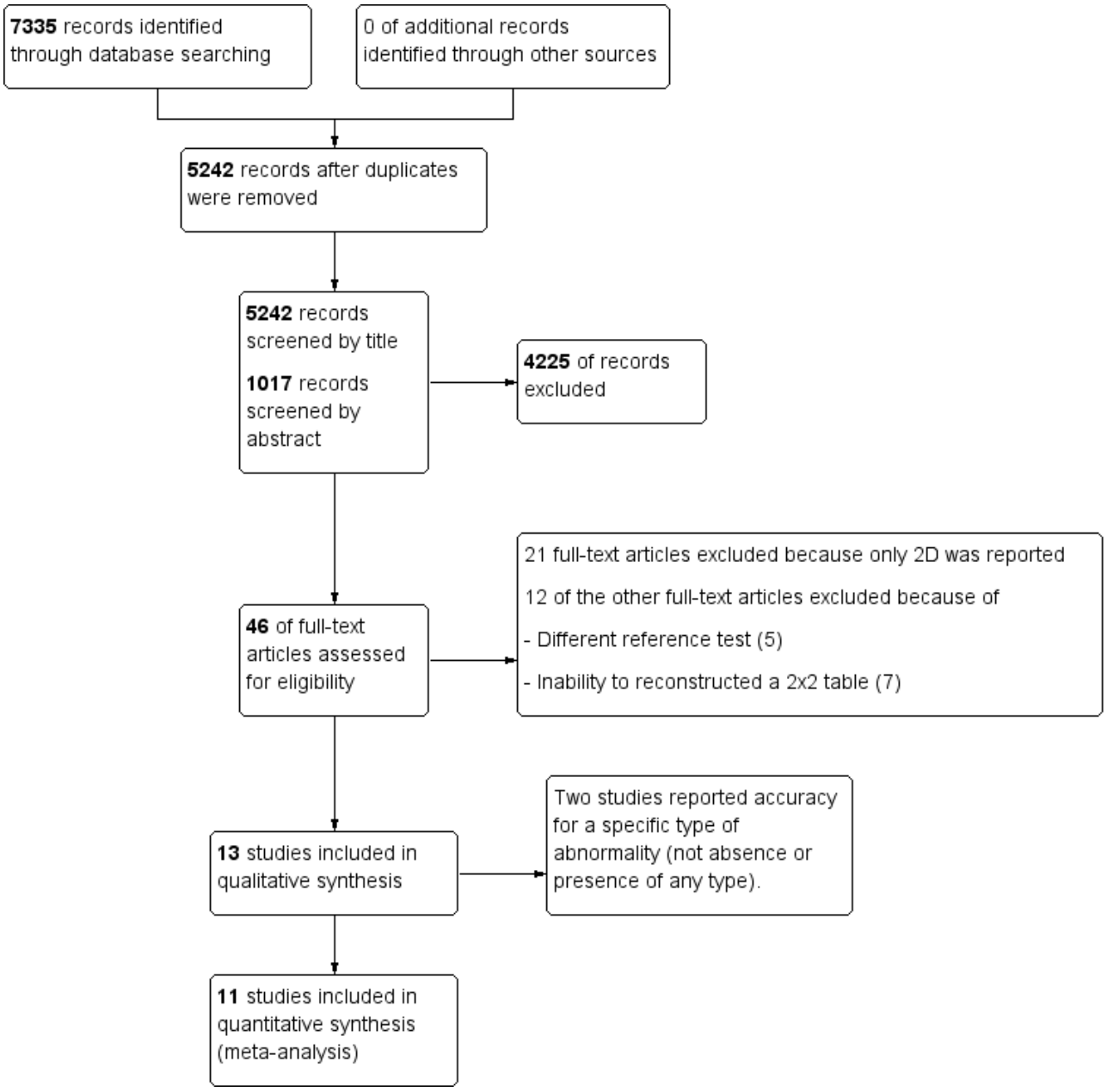
Results of the search

In total, we identified 7335 studies. After removing duplicates, two review authors (LLN and FJRH) independently screened 5242 studies by title. We carried out a handsearch by screening the reference lists of all included articles as well as available online conference abstracts in the field of gynaecology; this yielded no additional studies. Of the 5242 studies screened by title, we found that 1017 required screening by abstract. Both review authors independently screened 1017 abstracts. Subsequently, both review authors assessed 46 full-text articles for eligibility. We resolved disagreements through short discussion. For two studies, we consulted a third review author (HAMB) and reached agreement.

We found more than 10 studies reporting 3D SIS (compared with 2D SIS) and included all of them in the review. Because the accuracy of 3D SIS is our main question and we found enough studies, we did not include studies reporting only 2D SIS (21 studies). We excluded another 12 studies because 2×2 tables could not be constructed or because investigators used a different index test or reference test; five of those 12 studies (Abou-Salem 2010; Ahmad 2014; Khan 2011; Makris 2005; Makris 2007a) did not use hysteroscopy (but used histology) as the reference standard. Ayida 1996 compared 2D and 3D SIS in five participants without using a reference standard. Ayoubi 2002 reported about software used to create a virtual hysteroscopy, which investigators tested in five participants. Jurisic 2013 used a different index test - 3D multi-slice - and for Ahmadi 2013 and Lagana 2014, it was impossible to build a 2×2 table.

We included the remaining 13 studies for qualitative synthesis. Two of 13 studies (Kowalczyk 2012; Nieuwenhuis 2014) reported accuracy only for type of abnormality (secondary objective). We were able to include the other 11 studies in a meta-analysis. For details of the screening and selection process, see Figure 1. We did not find multiple papers of the same study. Included studies were published and performed between 1999 and 2015 in the following countries: Italy, Greece, England, Poland, the Netherlands, Egypt and Canada. None of these studies reported conflicts of interest. We found two ongoing studies (see Characteristics of ongoing studies).

Figure 1. Study flow diagram.



Methodological quality of included studies

The main problem involving the quality of included studies is insufficient reporting of study methods, resulting in unclear risk of bias for several of the quality domains assessed. Therefore, we considered the overall quality of the evidence as low.

Of the 13 included studies, 12 were prospective observational cohort studies and one was an RCT (Katsetos 2013). We found that methodological quality was often difficult to assess, as researchers did not clearly report the required information. We contacted study authors because we found conflicting results presented in the text and tables (leading to inability to complete 2×2 tables). Responses were not always helpful, but in all cases, we were able to prepare 2×2 tables.

Figure 2 and Figure 3 show bias and applicability for all included studies for each domain as scored by QUADAS-2. Some empty spaces are apparent in Figure 3, all for the index test 2D SIS. In these cases, investigators reported only 3D SIS, so quality assessment for 2D SIS was not applicable.

Figure 2. Risk of bias and applicability concerns graph: review authors' judgements about each QUADAS-2 domain presented as percentages across included studies.

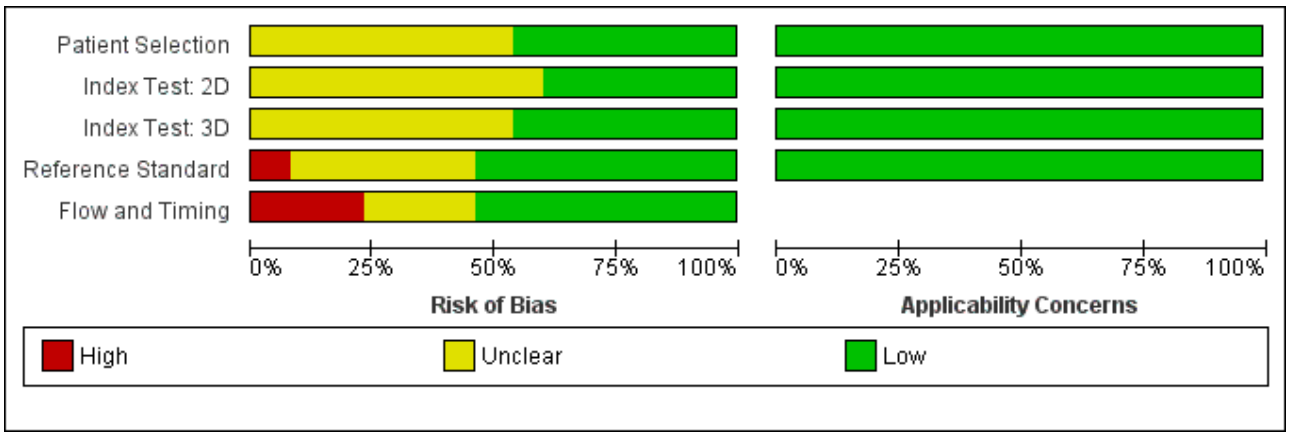
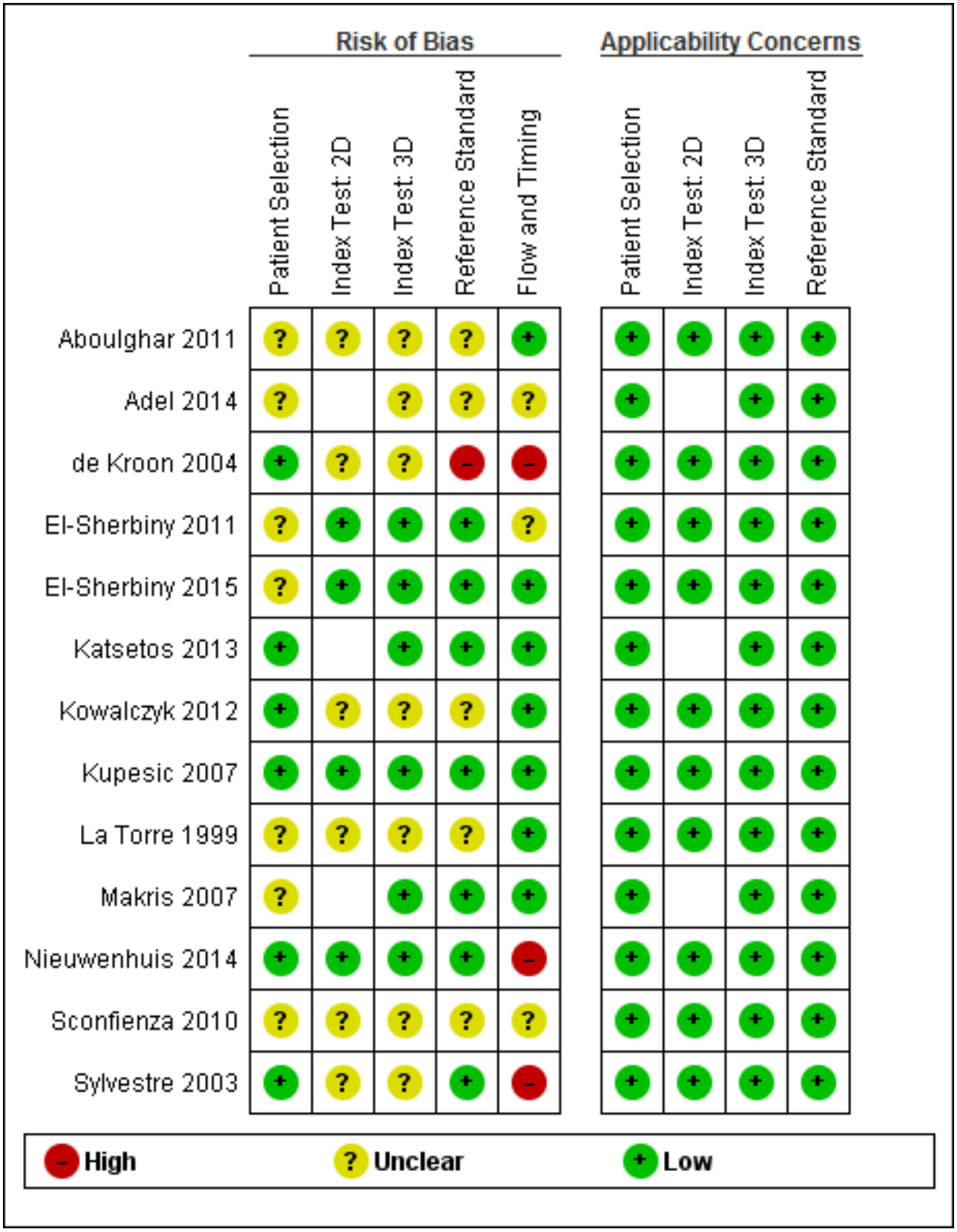


Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each QUADAS-2 domain for each included study.



For all studies, we had few applicability concerns. Figure 2 demonstrates that for all domains, half of the studies showed low risk of bias. We have addressed risk of bias for each domain separately below. The Characteristics of included studies section presents details of each domain separately for every study (Aboulghar 2011; Adel 2014; de Kroon 2004; El-Sherbiny 2011; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; Kupesic 2007; La Torre 1999; Makris 2007; Nieuwenhuis 2014; Sconfienza 2010; Sylvestre 2003).

- *Patient selection domain:* Six studies showed low risk of bias (de Kroon 2004; Katsetos 2013; Kowalczyk 2012; Kupesic 2007; Nieuwenhuis 2014; Sylvestre 2003), and the other seven showed unclear risk. Seven studies included participants with abnormal uterine bleeding (Adel 2014; de Kroon 2004; Katsetos 2013; Kowalczyk 2012; Makris 2007; Sconfienza 2010; Sylvestre 2003), three included participants with subfertility (Aboulghar 2011; El-Sherbiny 2011; Kupesic 2007) and three included both groups (El-Sherbiny 2015; La Torre 1999; Nieuwenhuis 2014). Seven studies did not state consecutive inclusion but did report the inclusion period. It remains unclear if included patients represented a random sample (unclear risk of bias). Six of 13 studies did not report a prior test (Adel 2014; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; Kupesic 2007; Sconfienza 2010). It is unclear whether these participants received 2D SIS and 3D SIS directly and 2D ultrasonography was passed over (uncommon in clinical practice). Therefore, risk of bias is unclear but cannot be considered high because included patients were clinically suspected of having intracavitary abnormalities. One study (El-Sherbiny 2011) included only patients with a negative prior test (normal 2D ultrasonography) and did not report patients with an abnormal prior test; therefore, we considered selection bias as unclear.
- *Index test domain:* All 13 studies reported 3D SIS (index test 1), 10 also reported 2D SIS (comparator test) and three had not studied 2D SIS (Adel 2014; Katsetos 2013; Makris 2007). Investigators performed all index tests for 2D SIS and 3D SIS in the same way across included studies and presented the same well-reported criteria for presence or absence of the target condition (a threshold was not applicable in all studies). We judged seven studies (Aboulghar 2011; Adel 2014; de Kroon 2004; Kowalczyk 2012; La Torre 1999; Sconfienza 2010; Sylvestre 2003) as having unclear risk of bias owing to unclear blinding. It was unclear whether index tests were interpreted without knowledge of the reference standard. We considered the others to have low risk of bias (n = 6).
- *Reference standard domain:* All studies used a reference standard (hysteroscopy) that was likely to correctly classify the condition. In five of 13 studies (Aboulghar 2011; Adel 2014; Kowalczyk 2012; La Torre 1999; Sconfienza 2010), it was unclear whether hysteroscopy was interpreted without knowledge of the index tests. One study (de Kroon 2004) did not blind observers to index test results; therefore, we judged this study to have high risk of bias.
- *Flow and timing domain:* We found most concerns regarding bias in this domain. We judged seven studies as having low risk of bias (Aboulghar 2011; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; Kupesic 2007; La Torre 1999; Makris 2007) and the other six as having unclear (Adel 2014; El-Sherbiny 2011; Sconfienza 2010) and high risk (de Kroon 2004; Nieuwenhuis 2014; Sylvestre 2003). We based this classification on several concerns. Six studies did not report failures and/or complications and did not always make clear whether all patients were included in the study and in the final analyses. Seven studies reported an unclear time interval between SIS and hysteroscopy (Aboulghar 2011; Adel 2014; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; La Torre 1999; Sylvestre 2003). Three studies did not always use the reference standard, resulting in high risk of verification bias. Most studies did not report experience in performing and interpreting index tests and reference standard results.

Twelve of 13 studies were observational cohort studies. Observational studies are the next best method after RCTs in terms of quality of evidence, according to the GRADE Working Group (Guyatt). Study design for diagnostic test studies is challenging. To address test accuracy, we can consider observational studies with a good design and performed in a proper manner according to the STARD checklist (Cohen 2015) as having higher quality. Although the design of included studies seems applicable, the main problem involving the quality of included studies is insufficient reporting of study methods, resulting in unclear risk of bias for the several QUADAS-2 domains. Therefore, we rated the quality of the evidence as low.

Findings

We found 13 studies that evaluated 3D SIS with hysteroscopy as a reference standard (Aboulghar 2011; Adel 2014; de Kroon 2004; El-Sherbiny 2011; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; Kupesic 2007; La Torre 1999; Makris 2007; Nieuwenhuis 2014; Sconfienza 2010; Sylvestre 2003); 10 of these also reported 2D SIS (Aboulghar 2011; de Kroon 2004; El-Sherbiny 2011; El-Sherbiny 2015; Kowalczyk 2012; Kupesic 2007; La Torre 1999; Nieuwenhuis 2014; Sconfienza 2010; Sylvestre 2003).

All 13 studies (1053 women) reported accuracy for 3D SIS, and 11 of these (846 women) reported accuracy in detecting the presence/absence of any abnormality (Aboulghar 2011; Adel 2014; de Kroon 2004; El-Sherbiny 2011; El-Sherbiny 2015; Katsetos 2013; Kupesic 2007; La Torre 1999; Makris 2007; Sconfienza 2010; Sylvestre 2003). Eight studies (Aboulghar 2011; Adel 2014; El-Sherbiny 2011; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; La Torre 1999; Nieuwenhuis 2014) reported the presence/absence of a specific abnormality (polyp or fibroid). Study size ranged from 23 to 180 participants. Prevalence of the target condition ranged from 14% to 96%.

Primary objectives

The accuracy of 3D SIS was based on the reporting accuracy of 11 studies in detecting intracavitary abnormalities (presence/absence), which we used for quantitative analyses. Summary estimates for sensitivity and specificity were 94.5% (95% confidence interval (CI) 90.6% to 96.9%) and 99.4% (95% CI 96.2% to 99.9%), respectively, evaluated against hysteroscopy.

Diagnostic accuracy of 3D SIS (index test 1) in comparison with 2D SIS

Meta-analysis of the eight studies (N = 716) directly comparing 2D SIS and 3D SIS (index test 1) (Aboulghar 2011; de Kroon 2004; El-Sherbiny 2011; El-Sherbiny 2015; Kupesic 2007; La Torre 1999; Sconfienza 2010; Sylvestre 2003) showed that both sensitivity and specificity are higher for 3D SIS than for 2D SIS (see Figure 4 for

forest plot, Figure 5 and Figure 6 for summary receiver operator characteristics (SROC) plots), although this difference was not statistically significant (P values of 0.07 for sensitivity and 0.10 for specificity). Figure 5 shows a SROC plot for both 2D SIS and 3D SIS. Figure 6 additionally shows sensitivity-specificity pairs from studies that studied both 2D SIS and 3D SIS; these are linked with a dashed line. Results for 2D SIS showed greater variation in specificity than those for 3D SIS. Mean sensitivity of 3D SIS was approximately the same as in the complete set of 11 3D SIS studies: sensitivity 96.9% (95% CI 91.9% to 98.8%); specificity 99.5% (95% CI 96.1% to 100%). Mean sensitivity for 2D SIS was 90.9% (95% CI 81.2% to 95.8%) and for specificity 96.3% (95% CI 86.1% to 99.1%). Inspection of the forest plots reveals that sensitivity and specificity for 3D SIS were similar among studies.

Figure 4. Forest plot of 2D SIS and 3D SIS; studies reporting abnormality or not.

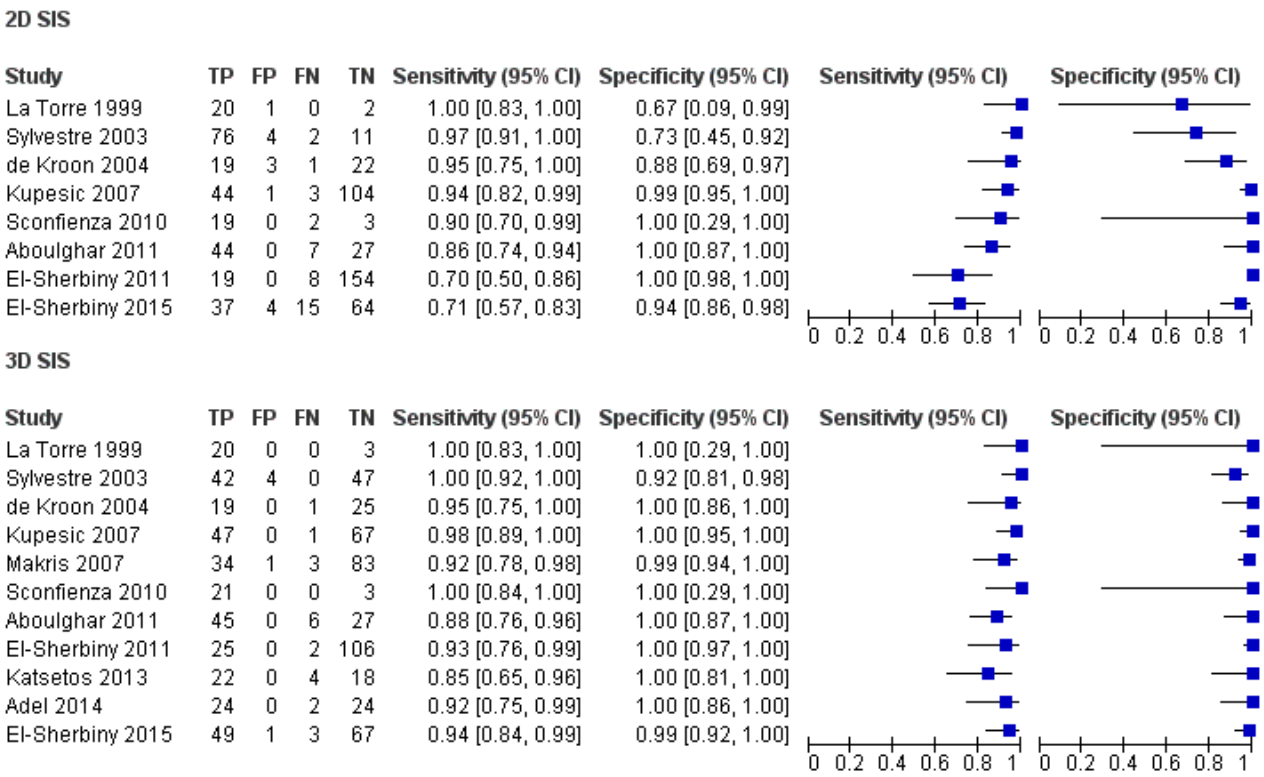


Figure 5. Summary ROC Plot of 2D SIS and 3D SIS.

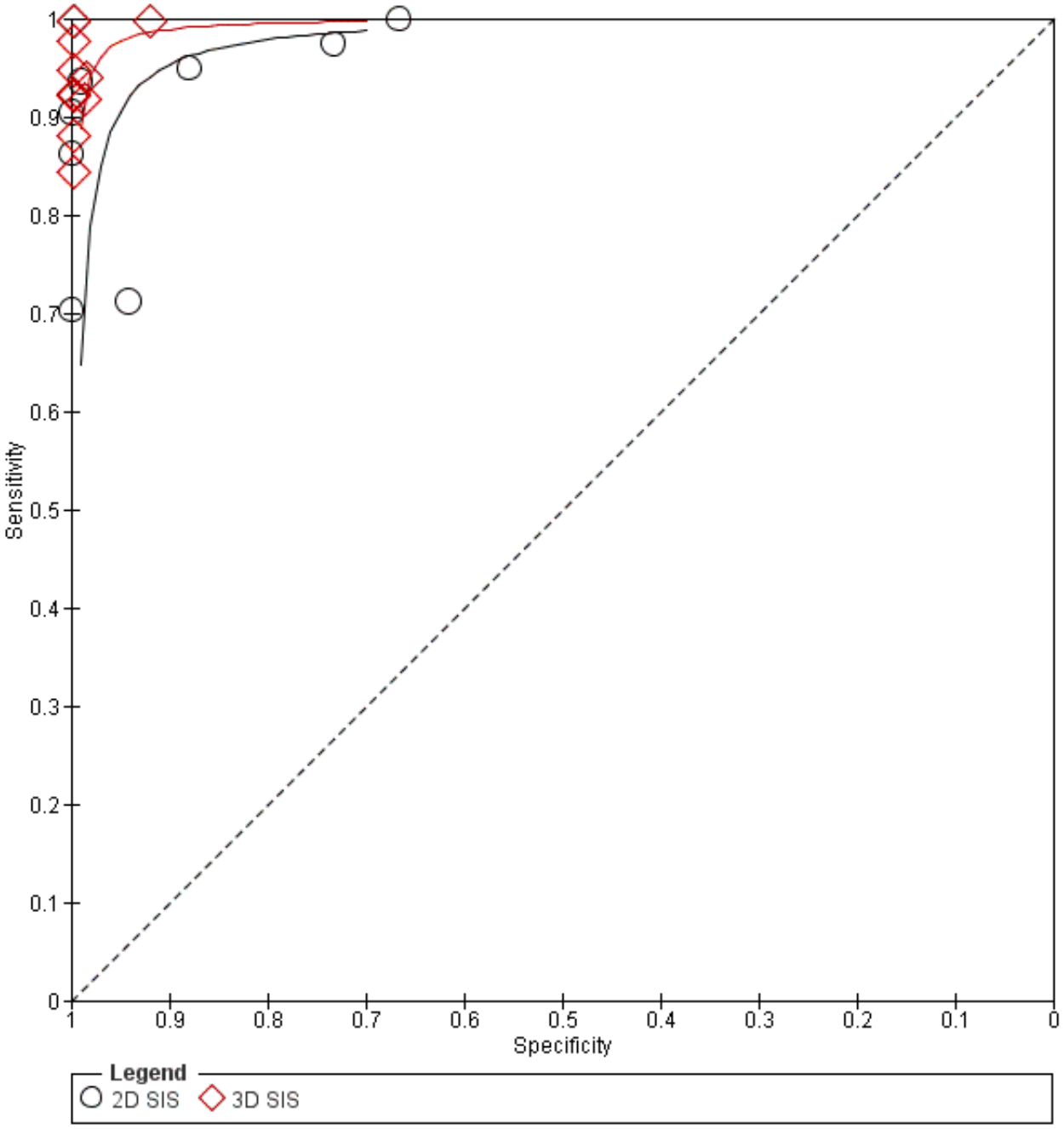
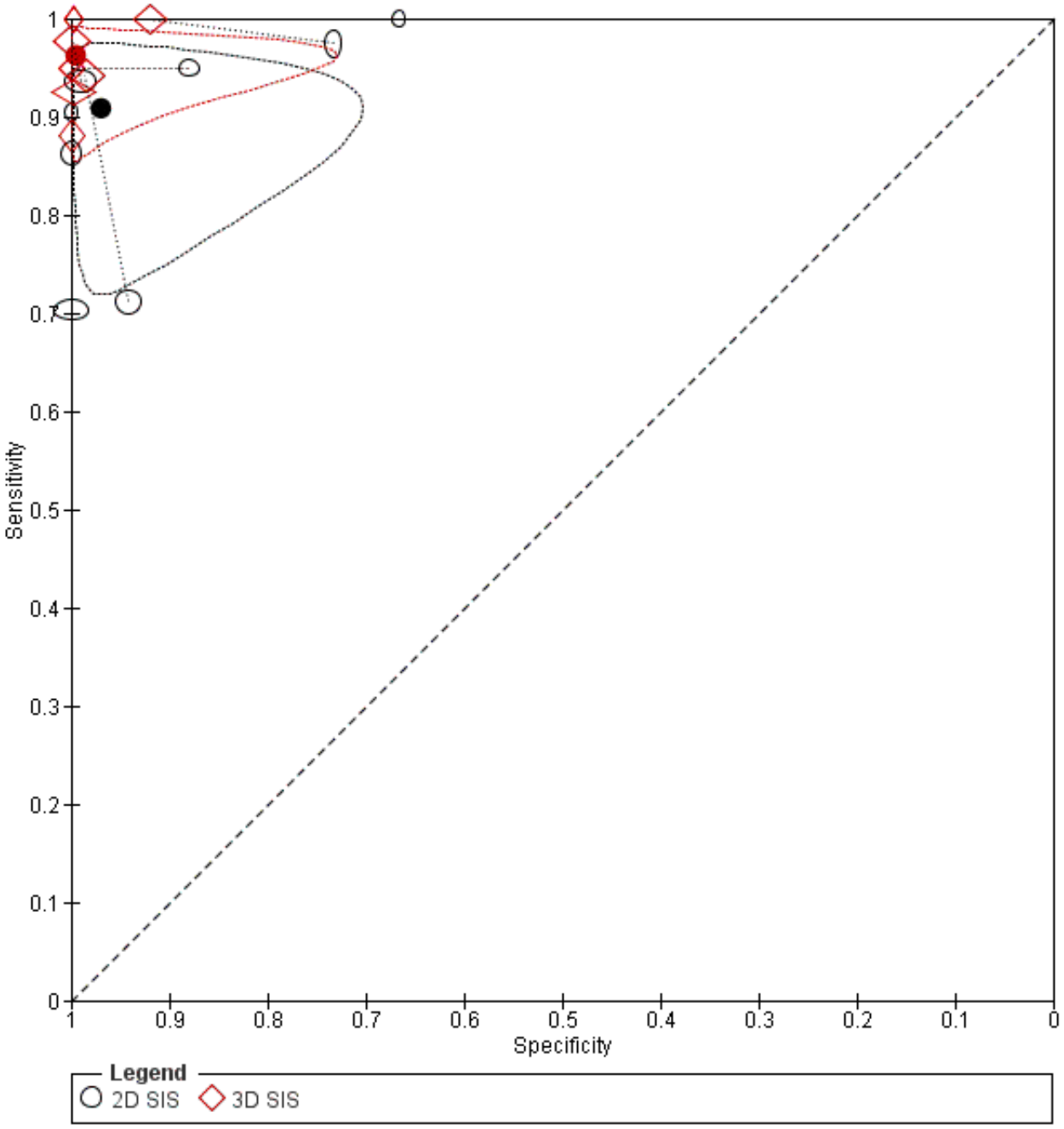


Figure 6. Summary ROC plot of 2D SIS and 3D SIS; sensitivity-specificity pairs from studies that studied both 2D SIS and 3D SIS are linked with a dashed line



Diagnostic accuracy of 2D SIS+3D SIS (index test 2) in comparison with 2D SIS

We found only one study (Nieuwenhuis 2014) that compared 2D SIS and 3D SIS together versus 2D SIS with hysteroscopy as a reference. Investigators did not report accuracy for the presence or absence of an abnormality but did report accuracy for detection of uterine polyps and submucous fibroids. We found only one study that used index test 2; therefore, additional (meta-)analyses were not possible.

To characterise the usefulness of the test in different prevalence scenarios, we calculated post-test probabilities (PPVs) for three different values of prevalence: 15%, 50% and 90%. PPV would be 96.0%, 99.3% and 99.9%, respectively.

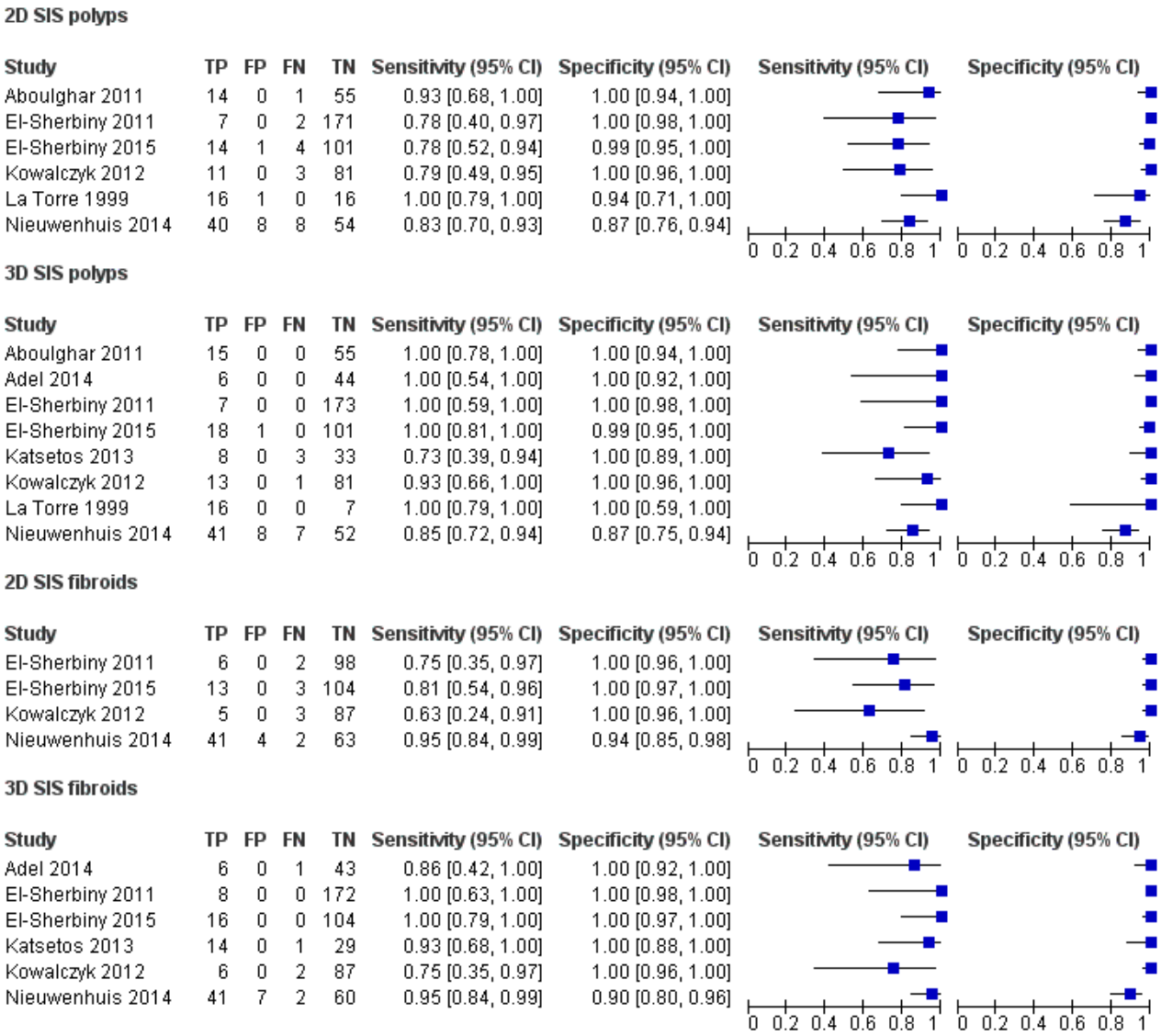
Secondary objectives

Diagnostic accuracy of 3D SIS in comparison with 2D SIS for type of abnormality: polyp or submucous fibroid

Polyps

Eight studies (Aboulghar 2011; Adel 2014; El-Sherbiny 2011; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; La Torre 1999; Nieuwenhuis 2014), with a total of 690 women, reported accuracy of 3D SIS in detecting polyps, and their summary sensitivity and specificity values were 96.3% (95% CI 79.4% to 99.4%) and 99.9% (95% CI 93.8% to 100%), respectively. This was comparable with the overall sensitivity for 3D SIS. Six studies reported accuracy for 2D SIS in detecting polyps; however, the models did not converge for this subgroup. Therefore, a formal comparison could not be made. However, five of six studies found that 3D SIS was more sensitive (improved approximately from 80% to 100%) and equally specific; one study found that 3D SIS was more specific and equally sensitive (Figure 7).

Figure 7. Forest plot of studies reporting type of abnormality (polyps and fibroids) for 2D SIS and 3D SIS.



Fibroids

Six studies (Adel 2014; El-Sherbiny 2011; El-Sherbiny 2015; Katsetos 2013; Kowalczyk 2012; Nieuwenhuis 2014), with a total of 599 women, reported accuracy of 3D SIS in detecting fibroids, and four studies reported accuracy of 2D SIS in detecting fibroids (Figure 7). Inspection of the forest plot reveals that in three of four studies, 3D SIS was more sensitive (60% to 90% to 75% to 100%) and equally specific. In the fourth study, 3D SIS was equally sensitive (95%) and less specific (94% to 90%). It is unclear if the higher sensitivity for 3D SIS is a result of more accurate measurement of protrusion with 3D SIS. Study authors did not report percentage of protrusion for 2D SIS nor for 3D SIS.

Diagnostic accuracy of 2D SIS+3D SIS in comparison with 2D SIS for type of abnormality: polyp or submucous fibroid

For detecting polyps, Nieuwenhuis 2014 reported that sensitivity and specificity improved from 83% and 87% for 2D SIS to 89% and 89% for 2D SIS+3D SIS.

For detecting fibroids, sensitivity improved with no effect on specificity: from 95% and 94% for 2D SIS to 98% and 94% for 2D SIS +3D SIS.

Sensitivity analyses

Original summary estimates for 3D SIS sensitivity and specificity were 94.5% and 99.4%, respectively. We found no missing or uninterpretable data. One study (de Kroon 2004) had high risk of bias for the reference standard. When we removed this study, summary sensitivity for 3D SIS was 94.6% (95% CI 90.3% to 97.1%) and summary specificity was 99.4% (95% CI 96.0% to 99.9%). Two of eight studies suffered from partial verification bias (de Kroon 2004; Sylvestre 2003). We were unable to perform sensitivity analyses on these two studies, as not enough data remained when they were removed. These were the oldest studies, and the forest plot shows that they had the highest estimates of sensitivity and the lowest estimates of specificity for 2D SIS. Removing them from the analyses might have lowered sensitivity and raised specificity values for 2D SIS. Three-dimensional SIS results were more uniform across different studies. Several studies had unclear risk of bias for patient selection and study sample. One of those studies (El-Sherbiny 2011) included patients with a normal cavity on 2D ultrasonography, which is different from all other studies in that investigators included patients with (complaints and) an abnormal 2D ultrasound. When we removed this study from the set, summary sensitivity for 3D SIS was 94.9% (95% CI 90.4% to 97.3%) and summary specificity was 99.1% (95% CI 95.2% to 99.8%).

DISCUSSION

Summary of main results

Thirteen studies (1053 women) reported accuracy of three-dimensional saline infusion sonography (3D SIS) and met the inclusion criteria. Summary estimates for sensitivity and specificity were 94.5% and 99.4%, respectively, evaluated against hysteroscopy. Meta-analysis (of eight studies) comparing two-dimensional (2D) SIS and 3D SIS showed no statistically significant differences (P values of 0.07 for sensitivity and 0.10 for specificity) in detecting intracavitary lesions with hysteroscopy as a reference standard. Mean sensitivity and mean specificity were higher for 3D SIS (96.9% and 99.5%) than for 2D SIS (90.9% and 96.3%). Detection of specific abnormalities (endometrial polyps and submucous fibroids) in most studies showed higher sensitivity for 3D SIS. Only one study reported on 2D SIS+3D SIS versus 2D SIS and found improved sensitivity (with equal or improved specificity). The design of included studies seems applicable. The main problem involving the quality of included studies is insufficient reporting of study methods, resulting in unclear risk of bias for several of the quality domains assessed. Therefore, we considered the overall quality of the evidence as low. Sensitivity analyses performed showed little effect on the data. We have provided a review summary in Summary of findings 1.

Strengths and weaknesses of the review

This is the first review conducted to study the accuracy of 3D SIS in detecting an intracavitary abnormality with hysteroscopy as a reference standard. We studied differences between 2D SIS and 3D SIS in meta-analysis and carried out subgroup analyses for polyps and fibroids. We conducted this review as published in the protocol and have reported minor differences below. We contacted the authors of potential studies for inclusion when important information was unclear, when only an abstract was available or when 2×2 tables could not be reconstructed. We kept to our strict inclusion criteria and included only studies in which a 2×2 table for 3D SIS (and 2D SIS) could be constructed. Another strength of the review is that we compared 2D SIS versus 3D SIS using only studies that reported both index tests. These within-study comparisons reporting both 2D SIS and 3D SIS provide results of comparisons of the same index tests and reference standard used in the same population. Differences in performance between index tests are not explained logically by different patients and settings. We analysed studies reporting only polyps or fibroids separately in the same manner.

However, this review has some limitations. In selecting studies, we chose to exclude studies in which pathology was the (only) reference standard. Studies using pathology as a reference focused on the origin of the abnormality (endometrium or myometrium) such that they included endometrial hyperplasia, which is usually seen as a diffuse endometrial lesion - not as a focal abnormality. In this review, we focused on the presence of an acquired, focal abnormality - not on the origin of the abnormality. We excluded studies reporting on uterine anomalies in general (not differentiating between global and focal intracavitary or congenital and acquired abnormalities) to keep groups of participants homogeneous. Lack of information on intrauterine hyperplasia and intrauterine adhesions limits the applicability of this review with respect to these populations because we can report only the accuracy of 3D SIS in detecting focal abnormalities.

For our secondary objective, we wanted to differentiate between polyps and fibroids. In the chosen analyses, we do not directly differentiate between polyps and fibroids but analyse polyp yes or no and fibroid yes or no. In an ideal situation, we wanted to directly differentiate between them, but we could not prepare a 2×2 table because several results are possible. Current sensitivity and specificity values for polyp yes or no and fibroid yes or no indicate how well we recognise that particular abnormality. When we prepared the protocol, we had hoped to find studies that used a second index test: comparison of combined results of 2D SIS and 3D SIS versus results of a comparator test (2D SIS). We found only one study; therefore we could perform no further analyses. These questions might be answered in the future.

We tried to contact all study authors to clarify various items on applied methods, but despite these efforts, a large number of QUADAS-2 (revised tool for quality assessment of diagnostic accuracy studies) items that needed to be answered to estimate methodological quality remain unclear. For example, several studies did not report failures and complications. This is notable in a consecutive patient sample (> 50 to 200 participants) for both SIS and hysteroscopy procedures. This information may show how successful SIS and hysteroscopy are, or it might suggest that missing data were not always reported, and that this could have influenced accuracy. If results from 2D SIS or 3D SIS are unclear, they should not be excluded from analyses but should be reported as positive. If no complications or failures are reported, one might think it is unclear whether all participants were included in the final analysis. This could have had a major impact on results, seen

mainly as an overestimation of accuracy for 2D SIS and 3D SIS. Another limitation of this review was the uncertainty of prior testing that could affect prevalence. On the other hand, prevalence is spread from 14% to 96% among included studies, and results are comparable; effects on study results might be considered minimal.

This review included a broad spectrum of patients and methodological differences that resulted in a heterogeneous group, impeding meta-analysis. For example, the broad spectrum resulted in a wide range of prevalence of the target condition. High prevalence suggests a highly selected patient population and can be seen as a weakness, although in clinical practice, it is not surprising that women with abnormal uterine bleeding (AUB) and a positive prior test will have a positive result at index testing. Even with this wide range of underlying disease prevalence, accuracy results remained similar for all studies.

Applicability of findings to the review question

All studies showed applicable patient selection and study design (at quality assessment using QUADAS-2), thus answering the question of diagnostic test accuracy. Findings of sensitivity and specificity observed between studies showed comparable results in terms of accuracy. Large variation is evident in prevalence of the target condition between studies. We included in the review both women with abnormal bleeding (premenopausal and postmenopausal) and women with subfertility. Prevalence was lowest in subfertile women and highest among women with AUB, as expected in clinical practice.

The time interval between SIS and hysteroscopy was unclear in some studies. This could have affected diagnosis or verification of the target condition. In clinical practice, these procedures follow each other quickly owing to availability, patient complaints and probability of diagnosis. It is not likely that time between procedures would be long enough for the target condition to change substantially. Another flaw that we found was unclear blinding of index results for the performer of the hysteroscopy (reference standard). Knowledge of the index test result could affect judgement of hysteroscopy findings, resulting in different assessment of test accuracy. In clinical practice, the physician performing hysteroscopy will usually know the findings of ultrasonography or SIS. During hysteroscopy, the shape, colour and texture of the cavity, endometrium and abnormality can be seen more clearly. Most likely, ultrasound findings have not affected the results of hysteroscopy because hysteroscopy probably overrules these findings. Therefore, we reported risk of bias as unclear or high, but we expect the effect on study results to be minimal.

A potential risk of verification bias results from the fact that the reference standard was not always performed when SIS was normal, leading patients to receive hysteroscopy only when SIS was abnormal. This may have influenced sensitivity and specificity. We do not know how this may have influenced these parameters, but we do know that El-Sherbiny 2011, which included only patients with a negative SIS, had comparable results for sensitivity and specificity.

Some studies did not report experience in performing and interpreting index tests. As 3D SIS is not a common practice in every hospital, it is likely that the procedure was performed by experts. Therefore, reported accuracy may be lower in everyday practice. As 2D SIS was probably also performed by experts, improved accuracy cannot be explained by the performer's experience. Mainly as a result of this poor reporting of methods in the included studies, we classified these uncertainties as having 'unclear risk of bias' in quality assessment based on QUADAS-2. It is possible that poor reporting/unclear risk of bias may have little effect on the results, but included studies should have reported the missing information. Therefore, we considered the overall quality of the evidence to be low. All studies proved to be applicable at quality assessment and can be considered representative for answering the main review question.

AUTHORS' CONCLUSIONS

Implications for practice

Low-quality evidence showed that 3D SIS is highly accurate in detecting intracavitary abnormalities. Meta-analysis revealed no statistically significant differences between 2D SIS and 3D SIS. Summary sensitivity and specificity are higher for 3D SIS, but margins of improvement are limited in that 2D SIS is already very accurate. 3D SIS is an alternative to 2D SIS when the technology and appropriate expertise are available. Both 2D SIS and 3D SIS should be considered alternatives to diagnostic hysteroscopy when intracavitary pathology is suspected in both subfertile women and those with abnormal uterine bleeding.

Implications for research

2D SIS and 3D SIS are frequently studied for detection of intracavitary abnormalities. Still, the sample size of studies in our meta-analysis might have been too small to show a difference. As the P value is close to 0.05, it would be interesting if more studies were conducted to determine whether a difference exists between 2D SIS and 3D SIS for detection of acquired focal uterine abnormalities. In addition, the low quality of evidence might suggest that solid methodological research must confirm previous findings.

It is unknown whether purchased materials (3D probe, software, etc.) and the learning curve outweigh any potential diagnostic advantage associated with 3D ultrasonic imaging. It is also unknown how effective 3D SIS is when 2D SIS is inconclusive. Future studies must evaluate these benefits and costs. Only if 3D SIS is found to be accurate and effective in everyday practice should widespread implementation and training be recommended. A well-powered randomised controlled trial comparing these tests in clinical practice and performing a cost-effectiveness analysis would be a good first step.

Finally, we were able to include in this review only four studies reporting both 2D SIS and 3D SIS for fibroids, even though this review focused on detection of abnormalities rather than on indications for each type of treatment. It would have been interesting to learn more about the accuracy of protrusion measurement. Current treatment of submucous fibroids is provided via operative hysteroscopy, but with upcoming alternative radiological interventions such as ultrasonically guided ablation, this additional morphological information may be crucial. Thus, future studies should focus on the classification of submucous fibroids.

ACKNOWLEDGEMENTS

We would like to thank the DTA Working Group and the Cochrane Gynaecology and Fertility Group for their specialist peer review comments. We thank Marian Showell (Information Specialist) for her great help with the search strategy. We thank Helen Nagels for her help and overall advice.

REFERENCES

References to studies included in this review

Aboulghar 2011 *{published data only}*

Aboulghar MM, Shoeir IK, Momtaz M, Mohammady ME, Ezzat H. A comparative study of 2-dimensional sonohysterography versus 3-dimensional sonohysterography in infertile patients with uterine cavity lesions and abnormalities. *Middle East Fertility Society Journal* 2011;**16**:67-71.

Adel 2014 *{published data only}*

Adel M, Kandil M, Abo-Elnasr M, Sanad Z, Farag H. Three-dimensional sonohysterography may replace hysteroscopy for women with perimenopausal bleeding. *Climacteric* 2014;**17**:55-9.

de Kroon 2004 *{published data only}*

de Kroon CD, Louwe LA, Trimbos JB, Jansen FW. The clinical value of 3-dimensional saline infusion sonography in addition to 2-dimensional saline infusion sonography in women with abnormal uterine bleeding: work in progress. *Journal of Ultrasound in Medicine* 2004;**23**:1433-40.

El-Sherbiny 2011 *{published data only}*

El-Sherbiny W, Nasr AS. Value of 3-dimensional sonohysterography in infertility work-up. *Journal of Minimally Invasive Gynecology* 2011;**18**:54-8.

El-Sherbiny 2015 *{published data only}*

El-Sherbiny W, El-Mazny A, Abou-Salem N, Mostafa WS. The diagnostic accuracy of two- vs three-dimensional sonohysterography for evaluation of the uterine cavity in the reproductive age. *Journal of Minimally Invasive Gynecology* 2015;**22**(1):127-31.

Katsetos 2013 *{published data only}*

Katsetos C, Radhakrishnan S, Koumousidis A, Kontoyannis M, Sanoulis V, Spaliaras D, et al. Comparison of transvaginal 3D sonohysterography with outpatient hysteroscopy in the evaluation of abnormal uterine bleeding. *Clinical & Experimental Obstetrics & Gynecology* 2013;**40**:74-7.

Kowalczyk 2012 *{published data only}*

Kowalczyk D, Guzikowski W, Wiecek J, Sioma-Markowska U. Clinical value of real time 3D sonohysterography and 2D sonohysterography in comparison to hysteroscopy with subsequent histopathological examination in perimenopausal women with abnormal uterine bleeding. *Neuroendocrinology Letters* 2012;**33**:212-6.

Kupesic 2007 *{published data only}*

Kupesic S, Plavsic BM. 2D and 3D hysterosalpingo-contrast-sonography in the assessment of uterine cavity and tubal patency. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2007;**133**:64-9.

La Torre 1999 *{published data only}*

La Torre R, De Felice C, De Angelis C, Coacci F, Mastrone M, Cosmi EV. Transvaginal sonographic evaluation of endometrial polyps: a comparison with two dimensional and three dimensional contrast sonography. *Clinical & Experimental Obstetrics & Gynecology* 1999;**26**:171-3.

Makris 2007 *{published data only}*

Makris N, Kalmantis K, Skartados N, Papadimitriou A, Mantzaris G, Antsaklis A. Three-dimensional hysterosonography versus hysteroscopy for the detection of intracavitary uterine abnormalities. *International Journal of Gynaecology & Obstetrics* 2007;**97**:6-9.

Nieuwenhuis 2014 *{published data only}*

Nieuwenhuis LL, de Vaate MA, Hehenkamp WJ, Heymans MW, van Baal MW, Brolmann HA, et al. Diagnostic and clinical value of 3D gel installation sonohysterography in addition to 2D gel installation sonohysterography in the assessment of intrauterine abnormalities. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2014;**175**:67-74.

Sconfienza 2010 *{published data only}*

Sconfienza LM, Lacelli F, Caldiera V, Perrone N, Piscopo F, Gandolfo N, et al. Three-dimensional sonohysterography for examination of the uterine cavity in women with abnormal uterine bleeding: preliminary findings. *Journal of Ultrasound* 2010;**13**:16-21.

Sylvestre 2003 *{published data only}*

Sylvestre C, Child TJ, Tulandi T, Tan SL. A prospective study to evaluate the efficacy of two- and three-dimensional sonohysterography in women with intrauterine lesions. *Fertility & Sterility* 2003;**79**:1222-5.

References to studies excluded from this review

Abou-Salem 2010 *{published data only}*

Abou-Salem N, Elmazny A, El-Sherbiny W. Value of 3-dimensional sonohysterography for detection of intrauterine lesions in women with abnormal uterine bleeding. *Journal of Minimally Invasive Gynecology* 2010;**17**:200-4.

Ahmad 2014 *{published data only}*

Ahmad RA, Sadek SM, Ragheb AS. Evaluation of the uterine cavity by magnetic resonance imaging, three dimensional hysterosonography and diagnostic hysteroscopy in women with pre- and post-menopausal bleeding. *Middle East Fertility Society Journal* 2015;**20**:70-8.

Ahmadi 2013 *{published data only}*

Ahmadi F, Rashidy Z, Haghighi H, Akhoond M, Niknejadi M, Hemat M, et al. Uterine cavity assessment in infertile women: sensitivity and specificity of three-dimensional hysterosonography versus hysteroscopy. *Iranian Journal of Reproductive Medicine* 2013;**11**:977-82.

Ayida 1996 *{published data only}*

Ayida G, Kennedy S, Barlow D, Chamberlain P. Contrast sonography for uterine cavity assessment: a comparison of conventional two-dimensional with three-dimensional

transvaginal ultrasound; a pilot study. *Fertility & Sterility* 1996;**66**:848-50.

Ayoubi 2002 *{published data only}*

Ayoubi JM, Fanchin R, Ferretti G, Pons JC, Bricault I. Three-dimensional ultrasonographic reconstruction of the uterine cavity: toward virtual hysteroscopy?. *European Radiology* 2002;**12**:2030-3.

Jurisic 2013 *{published data only}*

Jurisic A, Jurisic Z, Garalejic E, Arsic B, Jankovic-Raznatovic S, Arsic V. 3D multislice saline infusion sonography in preoperative evaluation of endometrial cavity. *Giornale Italiano di Ostetricia e Ginecologia* 2013;**35**:232-7.

Khan 2011 *{published data only}*

Khan F, Jamaat S, Al-Jaroudi D. Saline infusion sonohysterography versus hysteroscopy for uterine cavity evaluation. *Annals of Saudi Medicine* 2011;**31**:387-92.

Lagana 2014 *{published data only}*

Laganà AS, Ciancimino L, Mancuso A, Chiofalo B, Rizzo P, Triolo O. 3D sonohysterography vs hysteroscopy: a cross-sectional study for the evaluation of endouterine diseases. *Archives of Gynecology and Obstetrics* 2014;**290**(6):1173-8.

Makris 2005 *{published data only}*

Makris N, Kalmantis K, Skartados N, Xatzipappas J, Antsaklis A. Three-dimensional hysterosonography for the study of the endometrial cavity in infertile women: comparison with diagnostic hysteroscopy. *Gynecological Surgery* 2005;**2**:S42.

Makris 2007a *{published data only}*

Makris N, Skartados N, Kalmantis K, Mantzaris G, Papadimitriou A, Antsaklis A. Evaluation of abnormal uterine bleeding by transvaginal 3-D hysterosonography and diagnostic hysteroscopy. *European Journal of Gynaecological Oncology* 2007;**28**:39-42.

Mora-Guanche 2009 *{published data only}*

Mora-Guanche P, Sparacino L, Garcia-Guzman R, Bennett RA, Hernandez J, Palumbo A. Three-dimensional sonohysterography (SHG) has improved diagnostic accuracy for intrauterine pathology compared to two-dimensional SHG: a prospective pilot study in infertility patients. *Fertility & Sterility* 2009;**1**:S119.

Mora-Guanche 2010 *{published data only}*

Mora-Guanche P, Garcia-Guzman R, Bennett R, Iaconianni L, Hernandez J, Palumbo A. Diagnostic accuracy of three dimensional sonohysterography (SHG) for intrauterine abnormalities in infertility compared to two dimensional SHG. *Human Reproduction* 2010;**25**:i332.

References to ongoing studies

NCT02399501 *{published data only}*

Sonohysterography , 3D Ultrasonography and Hysteroscopy in Assessment of Uterine Factor in Cases of Female Infertility. Ongoing study March 2015; primary completion date January 2016 (final data collection date for primary outcome measure).

NCT02682433 *{published data only}*

D Sonohysterography vs Hysteroscopy: Study for the Evaluation of Intrauterine Abnormalities. Ongoing study April 2016; estimated study completion date January 2018.

Additional references

Abou Salem 2010

Abou-Salem N, Elmazny A, El-Sherbiny W. Value of 3-dimensional sonohysterography for detection of intrauterine lesions in women with abnormal uterine bleeding. *Journal of Minimally Invasive Gynecology* 2010;**17**(2):200-4.

Beemsterboer 2008

Beemsterboer SN, Thurkow AL, Verstraeten R, Brolmann HA. Reproducibility of saline contrast sonohysterography for the detection of intracavitary abnormalities in women with abnormal uterine bleeding. *Ultrasound in Obstetrics & Gynecology* 2008;**31**(4):445-9. [PUBMED: 18383471]

Betjes 2009

Betjes HE, Hanstede MM, Emanuel MH, Stewart EA. Hysteroscopic myomectomy and case volume hysteroscopic myomectomy performed by high- and low-volume surgeons. *The Journal of Reproductive Medicine* 2009;**54**(7):425-8. [PUBMED: 19691258]

Bij de Vaate 2010

Bij de Vaate AJ, Brolmann HA, van der Slikke JW, Emanuel MH, Huirne JA. Gel instillation sonohysterography (GIS) and saline contrast sonohysterography (SCSH): comparison of two diagnostic techniques. *Ultrasound in Obstetrics & Gynecology* 2010;**35**(4):486-9. [PUBMED: 20178106]

Bosteels 2015

Bosteels J, Kasius J, Weyers S, Broekmans FJ, Mol BW, D'Hooghe TM. Hysteroscopy for treating subfertility associated with suspected major uterine cavity abnormalities. *Cochrane Database of Systematic Reviews* 2015, Issue 2. [DOI: 10.1002/14651858.CD009461.pub3]

Cohen 2015

Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;**6**(11).

de Kroon 2003

de Kroon CD, de Bock GH, Dieben SW, Jansen FW. Saline contrast hysterosonography in abnormal uterine bleeding: a systematic review and meta-analysis. *BJOG* 2003;**110**(10):938-47. [PUBMED: 14550365]

Dijkhuizen 2003

Dijkhuizen FP, Mol BW, Bongers MY, Brolmann HA, Heintz AP. Cost-effectiveness of transvaginal sonography and saline infused sonography in the evaluation of menorrhagia. *International Journal of Gynaecology and Obstetrics* 2003;**83**(1):45-52. [PUBMED: 14511871]

Dongen 2011

van Dongen H, Timmermans A, Jacobi CE, Elskamp T, de Kroon CD, Jansen FW. Diagnostic hysteroscopy and saline infusion sonography in the diagnosis of intrauterine abnormalities: an assessment of patient preference. *Gynecological Surgery* 2011;**8**(1):65-70.

Emanuel 1995

Emanuel MH, Verdel MJ, Wamsteker K, Lammes FB. A prospective comparison of transvaginal ultrasonography and diagnostic hysteroscopy in the evaluation of patients with abnormal uterine bleeding: clinical implications. *American Journal of Obstetrics and Gynecology* 1995;**172**(2 Pt 1):547-52.

Fatemi HM 2010

Fatemi HM, Kasius JC, Timmermans A, van Disseldorp J, Fauser BC, Devroey P, et al. Prevalence of unsuspected uterine cavity abnormalities diagnosed by office hysteroscopy prior to in vitro fertilization. *Human Reproduction* 2010;**25**(8):1959-65. [PUBMED: PUBMED 20570971]

Gath 1987

Gath D, Osborn M, Bungay G, Iles S, Day A, Bond A. Psychiatric disorder and gynaecological symptoms in middle aged women: a community survey. *British Medical Journal (Clin Res Ed)* 1987 Jan 24;**294(6566)**:213-8. [PUBMED: PMID: 3101815]

Golan 2001

Golan A, Sagiv R, Berar M, Ginath S, Glezerman M. Bipolar electrical energy in physiologic solution - a revolution in operative hysteroscopy. *Journal of the American Association of Gynecologic Laparoscopists* 2001;**8**(2):252-8.

Guideline ACOG

ACOG technology assessment no. 8: Sonohysterography. *Obstetrics and Gynecology* 2012;**119**(6):1325. [DOI: 10.1097/AOG.0b013e31825af435.]

Guideline NVOG

Dr. W.J.K. Hehenkamp, gynaecoloog, NVOG (voorzitter). Heavy menstrual bleeding [Hevig menstrueel bloedverlies]. Guideline NVOG 2013:1-196. [http://nvog-documenten.nl/uploaded/docs/Richtlijn%20Hevig%20Menstrueel%20Bloedverlies%20ws%202013.pdf]

Guideline RCOG

National Institute for Health and Care Excellence (NICE). Heavy menstrual bleeding: assessment and management. NICE clinical guideline 44 2007, last updated: August 2016. [https://www.nice.org.uk/guidance/cg44/resources/heavy-menstrual-bleeding-assessment-and-management-975447024325]

Guideline SOGC

Singh S Best C Dunn S, Leyland N, Wolfman WL, Clinical Practice – Gynaecology Committee, Leyland N, Wolfman W, Allaire C, Awadalla A, Best C, Dunn S, Heywood M, Lemyre M, Marcoux V, Menard C, Potestio F, Rittenberg D, Singh S, Society of Obstetricians and Gynaecologists of Canada. Abnormal uterine bleeding in pre-menopausal women. *SOGC CLINICAL PRACTICE GUIDELINES* May 2013 (Replaces Aug 2001);**No. 292 (Replaces No. 106, Aug 2001)**:1-32. [https://sogc.org/wp-content/uploads/2013/01/106E-CPG-August2001.pdf]

Guyatt

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ.* 2008;**336**(7650):924-6.

Klatsky 2008

Klatsky PC, Tran ND, Caughey AB, Fujimoto VY. Fibroids and reproductive outcomes: a systematic literature review from conception to delivery. *American Journal of Obstetrics and Gynecology* 2008;**198**(4):357.

Kroon 2006

Kroon CD, Jansen FW. Saline infusion sonography in women with abnormal uterine bleeding: an update of recent findings. *Current Opinion in Obstetrics and Gynecology* 2006;**18**(6):653-7.

Lee 2006

Lee C, Salim R, Ofili-Yebovi D, Yazbek J, Davies A, Jurkovic D. Reproducibility of the measurement of submucous fibroid protrusion into the uterine cavity using three-dimensional saline contrast sonohysterography. *Ultrasound in Obstetrics & Gynecology* 2006;**28**(6):837-41. [PUBMED: 17019741]

Leeflang 2008

Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine* 2008;**149**(12):889-97. [PUBMED: 19075208]

Lethaby 2002

Lethaby A, Vollenhoven B, Sowter M. Efficacy of pre-operative gonadotrophin hormone releasing analogues for women with uterine fibroids undergoing hysterectomy or myomectomy: a systematic review. *BJOG.* 2002;**109**(10):108-9.

Liu 2007

Liu Z, Doan QV, Blumenthal P, Dubois RW. A systematic review evaluating health-related quality of life, work impairment, and health-care costs and utilization in abnormal uterine bleeding. *Value Health* 2007;**10**(3):183-9.

Mavrelos 2011

Mavrelos D, Naftalin J, Hoo W, Ben-Nagi J, Holland T, Jurkovic D. Preoperative assessment of submucous fibroids by three-dimensional saline contrast sonohysterography. *Ultrasound in Obstetrics & Gynecology* 2011;**38**(3):350-4. [PUBMED: 21584888]

Metello 2008

Metello J, Relva A, Milheras E, Colaco J, Retto H. Hysteroscopic diagnostic accuracy in post-menopausal bleeding. *Acta Medica Portuguesa* 2008;**21**(5):483-8.

Munro 2011

Munro MG, Critchley HO, Fraser IS, FIGO Menstrual Disorders Working Group. The FIGO classification of causes of abnormal uterine bleeding in the reproductive years. *Fertility & Sterility* 2011;**95**(7):2204-8.

Parsons 1993

Parsons AK, Lense JJ. Sonohysterography for endometrial abnormalities: preliminary results. *Journal of Clinical Ultrasound* 1993;**21**(2):87-95.

Patterson 1994

Patterson-Keels LM, Selvaggi SM, Haefner HK, Randolph JF Jr. Morphologic assessment of endometrium overlying submucosal leiomyomas. *Journal of Reproductive Medicine* 1994;**39**(8):579..

Pritts 2009

Pritts EA, Parker WH, Olive DL. Fibroids and infertility: an updated systematic review of the evidence. *Fertility & Sterility* 2009;**91**(4):1215..

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982-90. [PUBMED: 16168343]

Salim 2005

Salim R, Lee C, Davies A, Jolaoso B, Ofuasia E, Jurkovic D. A comparative study of three-dimensional saline infusion sonohysterography and diagnostic hysteroscopy for the classification of submucous fibroids. *Human Reproduction* 2005;**20**(1):253-7.

Sancho 2016

Sancho JM, Delgado VS, Valero MJ, Soteras MG, Amate VP, Carrascosa AA. Hysteroscopic myomectomy outcomes after 3-month treatment with either Ulipristal Acetate or GnRH analogues: a retrospective comparative study. *European Journal of Obstetric and Gynecologic Reproductive Biology* 2016;**198**:127-30.

Santer 2005

Santer M, Warner P, Wyke S. A Scottish postal survey suggested that the prevailing clinical preoccupation with heavy periods does not reflect the epidemiology of reported symptoms and problems. *Journal of Clinical Epidemiology* 2005;**58**(11):1206-10.

Stewart 1996

Stewart EA, Nowak RA. Leiomyoma-related bleeding: a classic hypothesis updated for the molecular era. *Human Reproduction Update* 1996;**2**(4):295.

Takwoingi 2015

Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Statistical Methods in Medical Research* 2015 Jun 26. [PUBMED: pii: 0962280215592269]

Tarneja 2002

Tarneja P, Dugal BS. Hysteroscopy: past, present and future. *Medical Journal of Armed Forces India* 2002;**58**(4):293-4.

Terry 2009

Terry S, Banks E, Harris K, Duvivier R, Dar P. Comparison of 3-dimensional with 2-dimensional saline infusion sonohysterograms for the evaluation of intrauterine abnormalities. *Journal of Clinical Ultrasound* 2009;**37**(5):258-62. [PUBMED: 19208422]

Tur-Kaspa 2006

Tur-Kaspa I, Gal M, Hartman M, Hartman J, Hartman A. A prospective evaluation of uterine abnormalities by saline infusion sonohysterography in 1,009 women with infertility or abnormal uterine bleeding. *Fertility & Sterility* 2006;**86**(6):1731-5. [PUBMED: 17007850]

van den Bosch 2008

Van den Bosch T, Verguts J, Daemen A, Gevaert O, Domali E, Claerhout F. Pain experienced during transvaginal ultrasound, saline contrast sonohysterography, hysteroscopy and office sampling: comparative study. *Ultrasound in Obstetrics & Gynecology* 2008;**31**:346-51.

van den Bosch 2015

Van den Bosch T, Ameye L, Van Schoubroeck D, Bourne T, Timmerman D. Intra-cavitary uterine pathology in women with abnormal uterine bleeding: a prospective study of 1220 women. *Facts, Views & Visions in Obgyn* 2015;**7**(1):17–24.

van den Bosch 2015a

Van den Bosch T, Dueholm M, Leone FP, Valentin L, Rasmussen CK, Votino A, et al. Terms, definitions and measurements to describe sonographic features of myometrium and uterine masses: a consensus opinion from the Morphological Uterus Sonographic Assessment (MUSA) group. *Ultrasound in Obstetrics & Gynecology* 2015;**46**(3):284-98.

van Dongen 2007

van Dongen H, de Kroon CD, Jacobi CE, Trimbos JB, Jansen FW. Diagnostic hysteroscopy in abnormal uterine bleeding: a systematic review and meta-analysis. *BJOG* 2007;**114**(6):664-75.

van Hanegem 2011

van Hanegem N, Breijer MC, Khan KS, Clark TJ, Burger MP, Mol BW, et al. Diagnostic evaluation of the endometrium in postmenopausal bleeding: an evidence-based approach. *Maturitas* 2011;**68**(2):155-64.

van Hanegem 2016

van Hanegem N, Breijer MC, Slockers SA, Zafarmand MH, Geomini P, Catshoek R, et al. Diagnostic workup for postmenopausal bleeding: a randomised controlled trial. *BJOG* 2017;**124**(2):231-40.

Wamsteker 1993

Wamsteker K, Emanuel MH, de Kruif JH. Transcervical hysteroscopic resection of submucous fibroids for abnormal uterine bleeding: results regarding the degree of intramural extension. *Obstetrics and Gynecology* 1993;**82**(5):736-40. [PUBMED: 8414318]

Werbrouck 2011

Werbrouck E, Veldman J, Luts J, Van Huffel S, Van Schoubroeck D, Timmerman D, et al. Detection of endometrial pathology using saline infusion sonography versus gel instillation sonography: a prospective cohort study. *Fertility & Sterility* 2011;**95**(1):285-8.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

Widrich 1996

Widrich T, Bradley LD, Mitchinson AR, Collins RL. Comparison of saline infusion sonography with office hysteroscopy for the evaluation of the endometrium. *American Journal of Obstetrics and Gynecology* 1996;**174**(4):1327-34.

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Aboulghar 2011

Study characteristics			
Patient sampling	Prospective cohort		
Patient characteristics and setting	Sample size: 70 Mean age: 30.8 Presentation: infertile patients with the assumption of an uterine abnormality on HSG or 2D ultrasonography Setting: outpatient clinic, Cairo University Hospital, Egypt		
Index tests	2D SIS and 3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, adhesions, anomalies Reference standard: hysteroscopy (and laparoscopy)		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Aboulghar 2011 *(Continued)*

		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Yes		
		Low	

Adel 2014

Study characteristics	
Patient sampling	Prospective cohort
Patient characteristics and setting	Sample size: 50 Mean age: 47.5 (SD 5.6) Presentation: perimenopausal bleeding; unclear whether a prior test was performed Setting: University Hospital, Egypt
Index tests	3D SIS
Target condition and reference standard(s)	Target condition: fibroids, polyps, blood, endometrial mass, thick endometrium

Adel 2014 *(Continued)*

	Reference standard: hysteroscopy		
Flow and timing			
Comparative			
Notes			

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	No		
Were withdrawals explained?	Unclear		
		Unclear	

de Kroon 2004

Study characteristics			
Patient sampling	Prospective cohort		
Patient characteristics and setting	Sample size: 49 Mean age: 42.3 Presentation: AUB (premenopausal and postmenopausal) in individual suspected of having intrauterine abnormalities on 2D ultrasound Setting: University Hospital Leiden, The Netherlands		
Index tests	2D SIS and 3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, endometrial folds Reference standard: hysteroscopy and histology		
Flow and timing			
Comparative			
Notes	Partial verification bias; individuals with negative findings on ultrasound did not receive hysteroscopy but received follow-up		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

de Kroon 2004 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
		High	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	No		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Yes		
		High	

El-Sherbiny 2011

Study characteristics			
Patient sampling	Prospective cohort		
Patient characteristics and setting	Sample size: 180 Mean age: 25.4 (2.8) Presentation: women attending infertility workup with a **normal** 2D ultrasound and HSG Setting: outpatient clinic, University Medical Center, Cairo, Egypt		
Index tests	2D SIS and 3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, adhesions, endometrial folds, anomalies Reference standard: hysteroscopy		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

El-Sherbiny 2011 *(Continued)*

DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	No		
Were withdrawals explained?	Unclear		
		Unclear	

El-Sherbiny 2015

Study characteristics	
Patient sampling	Comparative observational

El-Sherbiny 2015 *(Continued)*

Patient characteristics and setting	Sample size: 120 Mean age: 28.5 Presentation: infertility, menstrual disorders, recurrent pregnancy loss, clinically or ultrasonically suspected intrauterine abnormalities Setting: Hospital Cairo, Egypt
Index tests	2D SIS and 3D SIS
Target condition and reference standard(s)	Target condition: fibroid, polyp, adhesions, septum Reference standard: hysteroscopy
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**

El-Sherbiny 2015 *(Continued)*

DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Unclear		
		Low	

Katsetos 2013

Study characteristics			
Patient sampling	RCT		
Patient characteristics and setting	Sample size: 44 Mean age: 44.8 Presentation: recruitment from gynaecology and menstrual problem clinics; indications for referral were AUB and subfertility. Unclear whether a prior test was performed Setting: hospital, London, England		
Index tests	3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, abnormal endometrium, others Reference standard: hysteroscopy		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Yes		

Katsetos 2013 *(Continued)*

		Low	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Yes		
		Low	

Kowalczyk 2012

Study characteristics	
Patient sampling	Prospective cohort
Patient characteristics and setting	Sample size: 97 Mean age: 51 Presentation: perimenopausal bleeding; unclear whether a prior test was performed Setting: Medical University Hospital, hospitalised women, Poland
Index tests	2D SIS and 3D SIS
Target condition and reference standard(s)	Target condition: fibroid, polyp, endometrial folds

Kowalczyk 2012 *(Continued)*

	Reference standard? hysteroscopy		
Flow and timing			
Comparative			
Notes			

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	Yes		

Kowalczyk 2012 *(Continued)*

Were withdrawals explained?	Yes		
		Low	

Kupesic 2007

Study characteristics			
Patient sampling	Prospective cohort		
Patient characteristics and setting	Sample size: 2D SIS: 152; 3D SIS: 116 Mean age: 22 to 42 Presentation: subfertility or infertility; unclear whether a prior test was performed Setting: University Hospital, Zagreb, Croatia		
Index tests	2D SIS and 3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, adhesions, endometrial folds, anomalies Reference standard: hysteroscopy		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 3D			

Kupesic 2007 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	No		
Were withdrawals explained?	Unclear		
		Low	

La Torre 1999

Study characteristics	
Patient sampling	Prospective cohort
Patient characteristics and setting	Sample size: 23 Mean age: ? 35 to 45 years Presentation: combination of AUB and subfertility; patients were enrolled when a polyp was suspected on normal 2D ultrasound Setting: University Hospital, Rome, Italy
Index tests	2D SIS and 3D SIS
Target condition and reference standard(s)	Target condition: polyps, hyperplasia Reference standard: hysteroscopy
Flow and timing	

La Torre 1999 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Unclear		

La Torre 1999 *(Continued)*

		Low	

Makris 2007

Study characteristics			
Patient sampling	Prospective cohort		
Patient characteristics and setting	Sample size: 121 Mean age: 37 Presentation: patients with an abnormal 2D ultrasound or HSG were included Setting: University Hospital, Athens, Greece		
Index tests	3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, adhesions, endometrial folds, suspect for malignancy, anomalies Reference standard: hysteroscopy		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Makris 2007 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Yes		
		Low	

Nieuwenhuis 2014

Study characteristics			
Patient sampling	Prospective cohort		
Patient characteristics and setting	Sample size: 110 Mean age: 43.5 (SD 9.6) Presentation: AUB or subfertility; inclusion based on prior test (2D ultrasound) Setting: University Hospital, Amsterdam, The Netherlands		
Index tests	2D SIS, 3D SIS, 2D+3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, adhesions, endometrial folds, anomalies, other Reference standard: hysteroscopy (and histology)		
Flow and timing			
Comparative			
Notes	Partial verification bias; premenopausal women with negative SIS did not receive hysteroscopy		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

Nieuwenhuis 2014 *(Continued)*

DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	No		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Yes		
		High	

Sconfienza 2010

Study characteristics	
Patient sampling	Prospective cohort

Sconfienza 2010 *(Continued)*

Patient characteristics and setting	Sample size: 24 Mean age: ? Presentation: AUB; unclear whether a prior test was performed Setting: Radiology Department, Hospital, Italy
Index tests	2D SIS and 3D SIS
Target condition and reference standard(s)	Tartget condition: abnormality yes/no Reference standard? hysteroscopy
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			

Sconfienza 2010 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	Yes		
Were uninterpretable results or failures reported?	Unclear		
Were withdrawals explained?	Unclear		
		Unclear	

Sylvestre 2003

Study characteristics			
Patient sampling	Prospective cohort		
Patient characteristics and setting	Sample size: 93 Mean age: 35.7 (12) Presentation: subfertile women with abnormal 2D ultrasound or abnormal hysterosalpingogram Setting: Royal Victoria Teaching Hospital, Montreal, Canada		
Index tests	2D SIS and 3D SIS		
Target condition and reference standard(s)	Target condition: fibroid, polyp, endometrial folds, anomalies, myomas Reference standard? hysteroscopy Incorporation avoided? unclear		
Flow and timing			
Comparative			
Notes	Partial verification bias; women with negative SIS did not receive hysteroscopy		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Did the study avoid inappropriate exclusions?	Yes		
Was a consecutive sample of patients enrolled?	Yes		

Sylvestre 2003 *(Continued)*

		Low	**Low**
DOMAIN 2: Index Test 2D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test 3D			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Was partial verification avoided?	No		
Were uninterpretable results or failures reported?	Yes		
Were withdrawals explained?	Yes		
		High	

2D: two dimensional
3D: three dimensional
AUB: abnormal uterine blood loss
HSG: hysterosalpingography
SIS: saline infusion sonography

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abou-Salem 2010	Different reference standard (pathology)

Study	Reason for exclusion
Ahmad 2014	Different reference standard (pathology)
Ahmadi 2013	Inability to build a 2×2 table
Ayida 1996	Inability to build a 2×2 table; no reference standard used
Ayoubi 2002	Inability to build a 2×2 table; article reporting about virtual endoscopy
Jurisic 2013	Inability to build a 2×2 table; index test different; 3-dimensional saline infusion sonography (3D SIS) multi-slice
Khan 2011	Different reference standard (pathology)
Lagana 2014	Inability to build a 2×2 table
Makris 2005	Different reference standard (pathology)
Makris 2007a	Different reference standard (pathology)
Mora-Guanche 2009	Inability to build a 2×2 table (only abstract available)
Mora-Guanche 2010	Inability to build a 2×2 table (only abstract available)

Characteristics of ongoing studies *[ordered by study ID]*

NCT02399501

Trial name or title	Sonohysterography , 3D Ultrasonography and Hysteroscopy in Assessment of Uterine Factor in Cases of Female Infertility
Target condition and reference standard(s)	Target condition: uterine lesion Reference standard: diagnostic hysteroscopy
Index and comparator tests	Two dimensional ultrasound, three dimensional ultrasound, saline sonohysterography and hysteroscopy
Starting date	March 2015; primary completion date January 2016 (final data collection date for primary outcome measure)
Contact information	Ahmed Maged, Kasr Alainy medical school, Cairo University
Notes	

NCT02682433

Trial name or title	D Sonohysterography vs Hysteroscopy: Study for the Evaluation of Intrauterine Abnormalities
Target condition and reference standard(s)	Target condition: uterine diseases Reference standard: diagnostic hysteroscopy

NCT02682433 *(Continued)*

Index and comparator tests	2D SIS, 3D SIS
Starting date	April 2016; estimated study completion date January 2018
Contact information	Gad Malinger, professor. 052-4262123. gadm@tlvmc.gov.il Tel-Aviv Sourasky Medical Center
Notes	Estimated enrolment 250. Allocation: non-randomized

D A T A

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 2D SIS	8	716
2 3D SIS	11	846
3 2D SIS polyps	6	608
4 3D SIS polyps	8	690
5 2D SIS fibroids	4	431
6 3D SIS fibroids	6	599
7 2D+3D SIS polyps	1	123
8 2D+3D SIS fibroids	1	117

Test 1. 2D SIS.

Test 2. 3D SIS.

Test 3. 2D SIS polyps.

Test 4. 3D SIS polyps.

Test 5. 2D SIS fibroids.

Test 6. 3D SIS fibroids.

Test 7. 2D+3D SIS polyps.

Test 8. 2D+3D SIS fibroids.

ADDITIONAL TABLES

Table 1. Assessment of methodological quality; QUADAS-2 and additional questions

Domain	Patient selection	Index tests	Reference standard	Flow and timing
Description	Describe methods of patient selection Describe included participants (prior testing, presentation, setting)	Describe the index test and how it was conducted and interpreted	Describe the reference standard and how it was conducted and interpreted State the target condition	Describe any participants who did not receive the index test(s) and/or reference standard or who were excluded from the 2×2 table (refer to flow diagram) Describe the time interval and any interventions between index test(s) and reference standard
QUADAS-2 signalling questions (yes/no/unclear)	Was a consecutive or random sample of participants enrolled? Did the study avoid inappropriate exclusions?	Were the index test results interpreted without knowledge of results of the reference standard? If a threshold was used, was it prespecified?	Is the reference standard likely to correctly classify the target condition? Were the reference standard results interpreted without knowledge of results of the index test?	Was the interval between index test(s) and reference standard appropriate? Did all participants receive a reference standard? Did participants receive the same reference standard? Were all participants included in the analysis?
Additional signalling questions	How was participant recruitment arranged (based on presenting symptoms or results from a previous test)?	Was execution of the index test described sufficiently to permit replication of the test?	Was the target condition specified? Were complications with the reference standard registered?	Was partial verification bias avoided? Were uninterpretable results reported?

Table 1. Assessment of methodological quality; QUADAS-2 and additional questions *(Continued)*

		Were criteria for different index test findings well defined? Were complications with the index test registered? Is the amount of experience/training of the persons executing and reading the index tests specified?	Is the amount of experience/training of the persons executing and reading the reference test specified?	Were withdrawals explained?
Risk of bias (low/high/unclear)	Could the selection of participants have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct or its interpretation have introduced bias?	Could participant flow have introduced bias?
Concerns regarding applicability (low/high/unclear)	Are there concerns that included participants and setting do not match the review question?	Are there concerns that the index test, its conduct or its interpretation differs from the review question?	Is there concern that the target condition as defined by the reference standard does not match the review question?	

APPENDICES

Appendix 1. CENTRAL CRSO search strategy

Searched 01 March 2016

WEB platform

#1 MESH DESCRIPTOR Imaging, Three-Dimensional EXPLODE ALL TREES (928)
#2 ((Three-Dimensional and imag*)):TI,AB,KY (1511)
#3 ((3D and imag*)):TI,AB,KY (843)
#4 ((3 D and imag*)):TI,AB,KY (159)
#5 ((Three-Dimensional and sonogra*)):TI,AB,KY (30)
#6 ((3D and sonogra*)):TI,AB,KY (22)
#7 ((3 D and sonogra*)):TI,AB,KY (3)
#8 ((Three-Dimensional and sonohysterogra*)):TI,AB,KY (6)
#9 ((3D and sonohysterogra*)):TI,AB,KY (7)
#10 ((Three-Dimensional and SIS)):TI,AB,KY (2)
#11 ((3D and SIS)):TI,AB,KY (2)
#12 3dus:TI,AB,KY (8)
#13 (3 dus):TI,AB,KY (1)
#14 ((Three-Dimensional and ultraso*)):TI,AB,KY (312)
#15 ((3D and ultraso*)):TI,AB,KY (189)
#16 ((3 D and ultraso*)):TI,AB,KY (46)
#17 (3D US):TI,AB,KY (16)
#18 ((Three-Dimensional and hystero*)):TI,AB,KY (13)
#19 ((3D and hystero*)):TI,AB,KY (10)
#20 ((3 D and hystero*)):TI,AB,KY (2)
#21 (three dimension):TI,AB,KY (14)
#22 (3 dimension):TI,AB,KY (4)
#23 ((3d and multiplanar)):TI,AB,KY (14)
#24 ((3 dimension* and multiplanar)):TI,AB,KY (3)
#25 ((three dimension* and multiplanar)):TI,AB,KY (23)
#26 ((three dimensional or 3d or 3 d)):TI,AB,KY (4347)

#27 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 (4368)
#28 MESH DESCRIPTOR Ultrasonography EXPLODE ALL TREES (7718)
#29 MESH DESCRIPTOR Endosonography EXPLODE ALL TREES (271)
#30 #28 OR #29 (7718)
#31 1990 TO 2016:YR (777345)
#32 #30 AND #31 (7262)
#33 #27 OR #32 (11430)
#34 ((uter* adj2 abnormal*)):TI,AB,KY (89)
#35 ((abnormal vagina* bleeding)):TI,AB,KY (15)
#36 ((intrauter* adj2 patholog*)):TI,AB,KY (13)
#37 (intrauter* adj2 abnormal*):TI,AB,KY (16)
#38 (endometri* adj2 abnormal*):TI,AB,KY (47)
#39 (uter* adj2 anomal*):TI,AB,KY (22)
#40 (endometri* adj2 anomal*):TI,AB,KY (1)
#41 (intrauterine adj2 anomal*):TI,AB,KY (1)
#42 (dysfunctional uter* bleeding):TI,AB,KY (121)
#43 DUB:TI,AB,KY (28)
#44 (heavy menstrual bleed*):TI,AB,KY (101)
#45 (postmenopaus* adj2 bleed*):TI,AB,KY (42)
#46 (perimenopaus* adj2 bleed*):TI,AB,KY (2)
#47 MESH DESCRIPTOR Menorrhagia EXPLODE ALL TREES (261)
#48 MESH DESCRIPTOR Metrorrhagia EXPLODE ALL TREES (77)
#49 (uter* adj2 h?emorrhag*):TI,AB,KY (558)
#50 menorrhagi*:TI,AB,KY (531)
#51 metrorrhagi*:TI,AB,KY (201)
#52 (endometri* adj2 lesion*):TI,AB,KY (56)
#53 (endometri* adj2 adhesion*):TI,AB,KY (19)
#54 (uter* adj2 lesion*):TI,AB,KY (10)
#55 (uter* adj2 adhesion*):TI,AB,KY (15)
#56 (intrauter* adj2 adhesion*):TI,AB,KY (27)
#57 (ovar* adj2 adhesion*):TI,AB,KY (12)
#58 (intrauter* adj2 lesion*):TI,AB,KY (15)
#59 polyp*:TI,AB,KY (5147)
#60 endometrio*:TI,AB,KY (1214)
#61(adnexal mass*):TI,AB,KY (36)
#62 MESH DESCRIPTOR Adenomyosis EXPLODE ALL TREES (5)
#63 adenomyosis:TI,AB,KY (65)
#64 MESH DESCRIPTOR Leiomyoma EXPLODE ALL TREES (409)
#65 myoma*:TI,AB,KY (443)
#66 infertil*:TI,AB,KY (3628)
#67 subfertil*:TI,AB,KY (492)
#68 leiomyoma*:TI,AB,KY (511)
#69 fibroid*:TI,AB,KY (350)
#70 (arcuate uter*):TI,AB,KY (2)
#71 (endometri* adj2 thick*):TI,AB,KY (716)
#72 (uter* adj2 malformation*):TI,AB,KY (25)
#73 (bicornuate adj2 uterus):TI,AB,KY (2)
#74 (intracavity abnormal*):TI,AB,KY (1)
#75 (uter* adj2 contour):TI,AB,KY (0)
#76 (uter* adj3 sept*):TI,AB,KY (29)
#77 (endometri* adj2 atroph*):TI,AB,KY (54)
#78 (endometri* adj2 tumo?r*):TI,AB,KY (64)
#79 (uter* adj2 malignan*):TI,AB,KY (14)
#80 (uter* adj2 cancer*):TI,AB,KY (548)
#81 (endometri* adj2 malignan*):TI,AB,KY (15)
#82 (endometri* adj2 cancer*):TI,AB,KY (595)
#83 (ovar* adj2 malignan*):TI,AB,KY (35)
#84 (ovar* adj2 cancer*):TI,AB,KY (2465)
#85 (uter* adj2 disorder*):TI,AB,KY (6)
#86 (uter* adj2 disease*):TI,AB,KY (423)
#87 (endometri* adj2 neoplasm*):TI,AB,KY (317)

#88 (uter* adj2 neoplasm*):TI,AB,KY (1945)
#89 (uter* adj2 patholog*):TI,AB,KY (509)
#90 (endometr* adj2 patholog*):TI,AB,KY (518)
#91 MESH DESCRIPTOR Adenomyoma EXPLODE ALL TREES (3)
#92 Adenomyoma*:TI,AB,KY (8)
#93 fibroma*:TI,AB,KY (37)
#94 fibromyoma*:TI,AB,KY (11)
#95 MESH DESCRIPTOR Infertility, Female EXPLODE ALL TREES (1025)
#96 MESH DESCRIPTOR Endometriosis EXPLODE ALL TREES (505)
#97 MESH DESCRIPTOR Uterine Diseases EXPLODE ALL TREES (3708)
#98 MESH DESCRIPTOR Polyps EXPLODE ALL TREES (642)
#99 MESH DESCRIPTOR Endometrial Hyperplasia EXPLODE ALL TREES (106)
#100 (Endometri* adj3 Hyperplas*):TI,AB,KY (333)
#101 MESH DESCRIPTOR Ovarian Diseases EXPLODE ALL TREES (2425)
#102 (Ovar* adj2 Disease*):TI,AB,KY (639)
#103 (ovar* adj2 mass*):TI,AB,KY (20)
#104 (ovar* adj2 cyst*):TI,AB,KY (285)
#105 (ovar* adj2 tumo?r*):TI,AB,KY (139)
#106 MESH DESCRIPTOR Gynecology EXPLODE ALL TREES (105)
#107 Gyn?ecology:TI,AB,KY (1901)
#108 (uter* cavit*):TI,AB,KY (234)
#109 (endometrial cavity):TI,AB,KY (44)
#110 (intracav* lesion*):TI,AB,KY (2)
#111 (intracav* abnormal*):TI,AB,KY (4)
#112 (uter* adj2 volume*):TI,AB,KY (162)
#113 (ovar* adj2 volume*):TI,AB,KY (98)
#114 #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53 OR #54 OR #55 OR #56 OR #57 OR #58 OR #59 OR #60 OR #61 OR #62 OR #63 OR #64 OR #65 OR #66 OR #67 OR #68 OR #69 OR #70 OR #71 OR #72 OR #73 OR #74 OR #75 OR #76 OR #77 OR #78 OR #79 OR #80 OR #81 OR #82 OR #83 OR #84 OR #85 OR #86 OR #87 OR #88 OR #89 OR #90 OR #91 OR #92 OR #93 OR #94 OR #95 OR #96 OR #97 OR #98 OR #99 OR #100 OR #101 OR #102 OR #103 OR #104 OR #105 OR #106 OR #107 OR #108 OR #109 OR #110 OR #111 OR #112 OR #113 (20885)
#115 #33 AND #114 (326)

Appendix 2. Clinicaltrials.gov search strategy

Searched 01 March 2016

Intracavitary abnormality or sonohysterography

1 hit

Appendix 3. ICTRP search strategy

Searched 01 March 2016

Intracavitary abnormality or sonohysterography

19 hits

Appendix 4. PubMed search strategy

Searched 01 March 2016

((("ultrasonography" [Subheading] OR "Ultrasonography"[Mesh] OR "Hysterosalpingography"[Mesh] OR ((ultrasonograph*[tiab] OR ultrasound[tiab] OR sonograph*[tiab] OR sonohysterograph*[tiab] OR hysterosonograph*[tiab]) AND (transvagina*[tiab] OR saline[tiab] OR gel[tiab] OR infusion[tiab] OR instillation[tiab] OR contrast[tiab] OR enhanced[tiab])) OR sonohysterosalpingograph*[tiab] OR hysterosonosalpingograph*[tiab] OR hysterosalpingosonograph*[tiab] OR gis[tiab] OR sis[tiab] OR shsg[tiab] OR hssg[tiab] OR shg[tiab] OR HyCoSy[tiab] OR tvs[tiab])) AND ("Imaging, Three-Dimensional"[Mesh:NoExp] OR threedimension*[tiab] OR "3 d"[tiab] OR 3d[tiab] OR 3dus[tiab] OR "3 dus"[tiab] OR 3dimension*[tiab] OR ((three[tiab] OR 3[tiab]) AND dimension*[tiab]))) AND (menorrhagi*[tiab] OR metrorrhagi*[tiab] OR infertil*[tiab] OR polyp[tiab] OR polyps[tiab] OR ((bleeding*[tiab] OR hemorrhag*[tiab]) AND (vagina*[tiab] OR uterus[tiab] OR uterine[tiab] OR disorder*[tiab] OR postmenopausal[tiab] OR premenopausal[tiab])) OR ((abnormalit*[tiab] OR lesion*[tiab] OR disease[tiab] OR diseases[tiab] OR diseased[tiab] OR patholog*[tiab] OR disorder*[tiab]) AND (uterine[tiab] OR uterus[tiab] OR cavum[tiab] OR cavity[tiab] OR intracavit*[tiab] OR intrauterine[tiab] OR endometri*[tiab])) OR myoma*[tiab] OR leiomyoma*[tiab] OR adenomyoma*[tiab] OR (fibroid*[tiab] AND (uterus[tiab] OR uterine[tiab])) OR fibroma*[tiab] OR

fibromyoma*[tiab] OR "Leiomyoma"[Mesh] OR "Myoma"[Mesh:NoExp] OR "Uterine Hemorrhage"[Mesh] OR "Infertility, Female"[Mesh] OR "Polyps"[Mesh:NoExp])

756 hits

Appendix 5. MEDLINE search strategy

From inception until 01 March 2016

OVID platform

1 (uter$ adj2 abnormal$).tw. (2567)
2 abnormal vagina$ bleeding.tw. (478)
3 (intrauter$ adj2 abnormal$).tw. (291)
4 (endometri$ adj2 abnormal$).tw. (629)
5 (uterine adj2 anomal$).tw. (565)
6 (endometri$ adj2 anomal$).tw. (34)
7 (intrauterine adj2 anomal$).tw. (88)
8 (uterine adj2 anomal$).tw. (565)
9 abnormal uter$ bleeding.tw. (1505)
10 dysfunctional uter$ bleeding.tw. (800)
11 DUB.tw. (675)
12 heavy menstrual bleeding.tw. (482)
13 (postmenopaus$ adj2 bleed$).tw. (1022)
14 (perimenopaus$ adj2 bleed$).tw. (58)
15 exp uterine hemorrhage/ (18508)
16 uter$ haemorrhag$.tw. (136)
17 uter$ hemorrhag$.tw. (565)
18 menorrhagi$.tw. (2838)
19 metrorrhagi$.tw. (1007)
20 (endometri$ adj2 lesion$).tw. (2070)
21 (endometri$ adj2 adhesion$).tw. (320)
22 (uter$ adj2 lesion$).tw. (560)
23 (uter$ adj2 adhesion$).tw. (229)
24 (ovar$ adj2 adhesion$).tw. (177)
25 (intrauter$ adj2 adhesion$).tw. (348)
26 (intrauter$ adj2 lesion$).tw. (122)
27 polyp$.tw. (228328)
28 endometrio$.tw. (22928)
29 adnexal mass$.tw. (1959)
30 adenomyosis.tw. (1935)
31 exp leiomyoma/ or exp myoma/ (20495)
32 infertil$.tw. (46817)
33 subfertil$.tw. (3932)
34 myoma$.tw. (4946)
35 leiomyoma$.tw. (11503)
36 fibroid$.tw. (4746)
37 (septate adj2 uterus).tw. (357)
38 arcuate uter$.tw. (64)
39 (endometri$ adj2 thick$).tw. (2307)
40 (uter$ adj2 malformation$).tw. (636)
41 (bicornuate adj2 uterus).tw. (419)
42 intracavity abnormal$.tw. (2)
43 (uter$ adj2 contour).tw. (27)
44 (uter$ adj3 sept$).tw. (780)
45 endometri$ atroph$.tw. (208)
46 (endometri$ adj2 tumor$).tw. (1626)
47 (endometri$ adj2 tumour$).tw. (223)
48 ((uter$ adj2 malignan$) or (uter$ adj2 cancer$)).tw. (6004)
49 ((endometri$ adj2 malignan$) or (endometri$ adj2 cancer$)).tw. (13969)
50 ((ovar$ adj2 malignan$) or (ovar$ adj2 cancer$)).tw. (43226)
51 (uterus adj2 disorder$).tw. (4)
52 (uterine adj2 disorder$).tw. (167)

53 (uterus adj2 disease$).tw. (49)
54 (uterine adj2 disease$).tw. (594)
55 (endometri$ adj2 neoplasm$).tw. (254)
56 (uterine adj2 neoplasm$).tw. (377)
57 (uterus adj2 neoplasm$).tw. (16)
58 (uterine adj2 pathолog$).tw. (590)
59 (uterus adj2 pathолog$).tw. (62)
60 (endometri$ adj2 pathолog$).tw. (1045)
61 Adenomyoma/ (422)
62 Adenomyo$.tw. (2933)
63 fibroma.tw. (6129)
64 fibromyoma$.tw. (700)
65 Infertility, Female/ (25115)
66 exp endometriosis/ or exp uterine diseases/ (163912)
67 Polyps/ (9807)
68 Endometrial Hyperplasia/ (3143)
69 (Endometri$ adj3 Hyperplasia).tw. (3264)
70 exp Ovarian Diseases/ (96073)
71 (ovar$ adj2 mass$).tw. (1902)
72 (ovar$ adj2 tumo?r).tw. (6368)
73 (ovar$ adj2 cyst$).tw. (7209)
74 Gynecology/ (16898)
75 (gynecolog$ or gynaecolog$).tw. (75147)
76 uter$ cavit$.tw. (3334)
77 endometrial cavity.tw. (607)
78 intracav$ lesion$.tw. (49)
79 intracav$ abnormal$.tw. (27)
80 (uter$ adj2 volume).tw. (778)
81 (ovar$ adj2 volume$).tw. (961)
82 (intrauter$ adj2 pathолog$).tw. (240)
83 or/1-82 (630584)
84 Imaging, Three-Dimensional/ (50364)
85 (Three-Dimensional and imag$).tw. (36577)
86 (3D and imag$).tw. (31111)
87 (3 D and imag$).tw. (6624)
88 (Three-Dimensional and sonogra$).tw. (1090)
89 (3D and sonogra$).tw. (821)
90 (3 D and sonogra$).tw. (134)
91 3dus.tw. (152)
92 3 dus.tw. (14)
93 (Three-Dimensional and ultraso$).tw. (5390)
94 (3D and ultraso$).tw. (4037)
95 (3 D and ultraso$).tw. (1410)
96 3D US.tw. (348)
97 three dimension.tw. (439)
98 3 dimension.tw. (118)
99 (3d and multiplanar).tw. (843)
100 (3 dimension$ and multiplanar).tw. (183)
101 (three dimension$ and multiplanar).tw. (1112)
102 (three dimensional or 3d or 3 d).ti,ab,hw. and us.fs. (9851)
103 or/84-102 (95836)
104 Ultrasonography/ (64239)
105 limit 104 to yr="1990 -Current" (24736)
106 103 or 105 (119036)
107 (3D and hystero$).tw. (103)
108 (3 D and hystero$).tw. (18)
109 (Three-Dimensional and sonohysterogra$).tw. (33)
110 (3D and sonohysterogra$).tw. (26)
111 (3 D and sonohysterogra$).tw. (2)
112 (Three-Dimensional and SIS).tw. (59)
113 (3D and SIS).tw. (38)
114 (3 D and SIS).tw. (9)

115 (Three-Dimensional and hystero$).tw. (119)
116 (3d and hycosy).tw. (8)
117 (3 d and hycosy).tw. (2)
118 (three dimension$ and hycosy).tw. (11)
119 or/107-118 (252)
120 83 and 106 (3693)
121 119 or 120 (3797)
122 exp animals/ not humans.sh. (4191570)
123 121 not 122 (3658)

Appendix 6. Embase search strategy

From inception until 01 March 2016

OVID platform

1 (uter$ adj2 abnormal$).tw. (3744)
2 (intrauter$ adj2 abnormal$).tw. (425)
3 (endometri$ adj2 abnormal$).tw. (856)
4 (uterine adj2 anomalies).tw. (713)
5 (endometri$ adj2 anomal$).tw. (47)
6 (intrauterine adj2 anomal$).tw. (130)
7 (uterine adj2 anomal$).tw. (935)
8 abnormal uter$ bleeding.tw. (2303)
9 dysfunctional uter$ bleeding.tw. (1024)
10 abnormal vagina$ bleeding.tw. (605)
11 DUB.tw. (903)
12 heavy menstrual bleeding.tw. (802)
13 intrauterine patholog$.tw. (321)
14 adnexal mass$.tw. (2850)
15 (uter$ adj2 malformation$).tw. (897)
16 exp uterus bleeding/ or uterine body disease/ (8339)
17 exp menorrhagia/ or "menorrhagia and metrorrhagia"/ (7890)
18 uter$ haemorrhag$.tw. (100)
19 uter$ hemorrhag$.tw. (535)
20 menorrhagi$.tw. (4288)
21 metrorrhagi$.tw. (1103)
22 (endometri$ adj2 lesion$).tw. (3052)
23 (endometri$ adj2 adhesion$).tw. (486)
24 (uter$ adj2 lesion$).tw. (723)
25 (uter$ adj2 adhesion$).tw. (322)
26 (intrauter$ adj2 lesion$).tw. (180)
27 (intrauter$ adj2 adhesion$).tw. (522)
28 polyp$.tw. (254616)
29 arcuate uter$.tw. (142)
30 endometrio$.tw. (32437)
31 endometrial cavity.tw. (896)
32 (endometri$ adj2 chang$).tw. (1164)
33 exp leiomyoma/ (15057)
34 exp uterus myoma/ or exp myoma/ (13856)
35 infertil$.tw. (62858)
36 subfertil$.tw. (5079)
37 myoma$.tw. (6767)
38 fibroid$.tw. (7555)
39 leiomyoma$.tw. (14056)
40 (septate adj2 uterus).tw. (561)
41 (bicornuate adj2 uterus).tw. (576)
42 intracavity abnormal$.tw. (2)
43 (uter$ adj2 contour).tw. (49)
44 (uter$ adj3 sept$).tw. (1227)
45 endometri$ atroph$.tw. (260)
46 (endometri$ adj2 tumor$).tw. (2126)
47 (endometri$ adj2 tumour$).tw. (316)

48 ((uter$ adj2 malignan$) or (uter$ adj2 cancer$)).tw. (7271)
49 ((endometri$ adj2 malignan$) or (endometri$ adj2 cancer$)).tw. (19880)
50 ((ovar$ adj2 malignan$) or (ovar$ adj2 cancer$)).tw. (58966)
51 (uterus adj2 disorder$).tw. (10)
52 (uterine adj2 disorder$).tw. (232)
53 (uterus adj2 disease$).tw. (78)
54 (uterine adj2 disease$).tw. (771)
55 (endometri$ adj2 neoplasm$).tw. (291)
56 (uterine adj2 neoplasm$).tw. (415)
57 (uterus adj2 neoplasm$).tw. (19)
58 (uterine adj2 patholog$).tw. (930)
59 (uterus adj2 patholog$).tw. (94)
60 (endometri$ adj2 patholog$).tw. (1561)
61 (endometri$ adj2 thick$).tw. (3739)
62 exp adenomyoma/ (568)
63 Adenomyo$.tw. (4050)
64 fibroma.tw. (6088)
65 fibromyoma$.tw. (597)
66 exp female infertility/ (37958)
67 endometriosis/ (28638)
68 exp uterus disease/ (204087)
69 exp endometrium tumor/ (46314)
70 polyp/ or endometrium polyp/ (17532)
71 endometrium hyperplasia/ (6235)
72 Endometrial Hyperplasia.tw. (3390)
73 exp ovary disease/ (158592)
74 (ovar$ adj2 mass$).tw. (2718)
75 (ovar$ adj2 tumo?r).tw. (8093)
76 uter$ cavit$.tw. (4628)
77 intracavity lesion$.tw. (7)
78 (postmenopaus$ adj2 bleed$).tw. (1464)
79 (perimenopaus$ adj2 bleed$).tw. (88)
80 (uter$ adj2 volume).tw. (1038)
81 (ovar$ adj2 volume$).tw. (1337)
82 gyn?ecology.tw. (39192)
83 or/1-82 (736804)
84 three dimensional imaging/ (68484)
85 limit 84 to yr="1990 -Current" (68477)
86 (Three-Dimensional adj2 imag$).tw. (8841)
87 (3 D adj2 imag$).tw. (2326)
88 (Three-Dimensional adj2 sonogra$).tw. (448)
89 (3D adj2 sonogra$).tw. (447)
90 (3 D adj2 sonogra$).tw. (46)
91 3dus.tw. (222)
92 3 dus.tw. (19)
93 (Three-Dimensional adj2 ultraso$).tw. (2938)
94 (3D adj2 ultraso$).tw. (2977)
95 3D US.tw. (564)
96 (3d adj2 multiplanar).tw. (221)
97 (3 dimension$ adj2 multiplanar).tw. (32)
98 (three dimension$ adj2 multiplanar).tw. (214)
99 or/85-98 (76270)
100 (Three-Dimensional and sonohysterogra$).tw. (56)
101 (3D and sonohysterogra$).tw. (60)
102 (3 D and sonohysterogra$).tw. (9)
103 (Three-Dimensional and SIS).tw. (74)
104 (3D adj2 SIS).tw. (16)
105 (3 D and SIS).tw. (9)
106 (Three-Dimensional adj2 hystero$).tw. (35)
107 (3D adj2 hystero$).tw. (37)
108 (3 D and hystero$).tw. (41)
109 (3d and hycosy).tw. (23)

110 (3 d and hycosy).tw. (2)
111 (three dimension$ and hycosy).tw. (23)
112 echography/ (248347)
113 Three dimensional.tw. (136027)
114 3dus.tw. (222)
115 3D.tw. (115055)
116 113 or 114 or 115 (219884)
117 112 and 116 (4707)
118 99 or 117 (77727)
119 83 and 118 (2423)
120 or/100-111 (257)
121 119 or 120 (2547)

Appendix 7. Gynaecology and Fertility Specialised Register for RCTs

From inception until 01 March 2016

PROCITE PLATFORM

Keywords CONTAINS "3D hysterosonography"or"3D sonography"or "3D transvaginal ultrasound"or "3D ultrasound" or "three dimensional"or Title CONTAINS "3D hysterosonography"or"3D sonography"or "3D transvaginal ultrasound"or "3D ultrasound" or "three dimensional"

12 hits

Appendix 8. Gynaecology and Fertility DTA Register

From inception until 01 March 2016

PROCITE PLATFORM

Title CONTAINS "three dimensional" or "3d" or "3 dimensional" or Keywords CONTAINS "three dimensional" or "3d" or "3 dimensional"

16 hits

CONTRIBUTIONS OF AUTHORS

LL Nieuwenhuis: conceiving of the review, coordinating the review, providing a clinical perspective, designing search strategies, providing a methodological perspective, writing the review.

FJR Hermans: providing a methodological perspective, writing the protocol.

AJM Bij de Vaate: conceiving of the protocol, providing general advice on the protocol.

MMG Leeflang: providing a methodological perspective, providing general advice on the review.

HAM Brölmann: providing a clinical perspective, providing a policy perspective, providing a consumer perspective, providing general advice on the review, performing previous work that was the foundation of the current review.

WJK Hehenkamp: providing a clinical perspective, providing a policy perspective, providing a consumer perspective, providing general advice on the review, performing previous work that was the foundation of the current review.

BWJ Mol: providing a clinical perspective, providing a policy perspective, providing a consumer perspective, providing general advice on the review, performing previous work that was the foundation of the current review.

TJ Clark: providing a clinical perspective, providing a policy perspective, providing a consumer perspective, providing general advice on the review, performing previous work that was the foundation of the current review.

JAF Huirne: conceiving of the review, providing a clinical perspective, providing a policy perspective, providing a consumer perspective, providing general advice on the review, performing previous work that was the foundation of the current review.

DECLARATIONS OF INTEREST

The review authors have no conflicts of interest and no financial ties to disclose.

SOURCES OF SUPPORT

Internal sources

- VU University Medical Center, Netherlands.

External sources

- None, Other.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Primary objectives

Review: The primary objective of this review was to evaluate the diagnostic accuracy of 3D SIS (index test 1) in comparison with 2D SIS in the diagnosis of *focally growing lesions (presence or not)* in women with AUB or subfertility with hysteroscopy used as the reference test.

Protocol: focally growing lesions (polyps or fibroids).

In the protocol, the term 'polyps' or 'fibroids' is used to explain what is meant by focally growing lesions. It might be confusing because in the secondary objective, we differentiate between them. To make the primary objective more clear, we chose presence or not.

Types of studies

We state the following: All diagnostic test accuracy studies, randomised controlled trials and prospective cohort studies for which a 2×2 contingency table could be reproduced in which 2D SIS and 3D SIS were evaluated with results of hysteroscopy as the reference standard were eligible for inclusion in the review.

Review: 2×2 contingency table could be reproduced.

Protocol: sufficient methodological quality.

As inclusion based on "sufficient methodological quality" turned out to be difficult to objectively assess, we included only studies for which a 2×2 table could be reproduced.

Review: Prospective cohort studies will be included.

Protocol: Prospective cohort studies will be included if enrolment was performed consecutively.

As enrolment is a QUADAS-2 item, we removed 'consecutively' as this is not an exclusion criterion.

Review: Second, all studies in which 2D SIS or 3D SIS alone was evaluated were considered eligible for inclusion, although if enough studies (10) were found, we preferred studies that reported both 2D SIS and 3D SIS.

Protocol: In the protocol, we stated only 'enough' studies. In the review, we specified this as 10 studies.

Review: added the following: 'Authors of unpublished studies (only congress abstract or published protocol available) were contacted to facilitate inclusion of additional useful data. Language restrictions were not applied'.

Protocol: did not provide above information. As this information is important for the search, we added it to the review.

Index tests

Review: added the following: 'SIS was defined as positive for a focal intrauterine lesion when any distortion of the endometrial lining was visualised (see below). Time between index test and reference standard should preferably be less than one month'.

Protocol: A positive test result was not yet defined; therefore, we added this to the review. Second, a time interval was not yet described.

Target condition

The review extended ultrasonographic features of (submucous) fibroids with a more recent and widely supported reference. We also added information about classification of submucous fibroids.

Review: Smoothly margined echogenic masses with a homogenous texture were classified as polyps (Parsons 1993), and a uterine fibroid was seen as a well-defined round lesion within the myometrium or attached to it, often showing shadows at the edge of the lesion and/or internal fan-shaped shadowing (van den Bosch). Typically, echogenicity varies and some hyperechogenicity may be present internally. Submucous fibroids were classified into fibroid types (types 0 to 2) using the FIGO PALM-COEIN classification for abnormal uterine bleeding:

type 0 = pedunculated intracavitary, 1 = submucosal < 50% intramural, 2 = submucosal ≥ 50% intramural (Munro 2011). Hyperplasia, adhesions and congenital anomalies were not considered as focally growing lesions.

Protocol: Smoothly margined echogenic masses with a homogenous texture are classified as polyps, and structures of mixed echogenicity disrupting endometrial continuity are described as submucous fibroids.

Reference standards

Review: added the following: 'The absence of an intracavitary abnormality was the clear vision of the entire cavity without any disruption of the endometrial lining'.

Protocol: Above information was lacking in the protocol and therefore was added to the review.

Electronic searches

Review: Performed search was presented in a more structured way.

Review: 'from inception till March 1, 2016'.

Protocol: 'from inception until present'.

Search of other resources

Review: Handsearch was carried out by screening the reference lists of all included articles.

Protocol: We will handsearch the reference lists of all included articles and will *contact experts in the field to obtain additional data*.

We did not contact any experts in the field other than listed coauthors and authors of included and (some) excluded studies.

Selection of studies

Review: added the following: 'ENDNOTE was used as a bibliographic management system. Duplicates were removed after each study was checked by hand. All studies were verified on multiple papers of the same study'.

Protocol: Above information was lacking in the protocol and therefore was added to the review.

Assessment of methodological quality

Review: added the following: 'The QUADAS-2 tool consists of four different domains: patient selection, index test, reference standard, flow and timing. Each domain includes questions used to assess risk of bias and to address applicability concerns'. 'To assess quality, all four domains were scored low, unclear or high risk of bias and low or high concern regarding applicability'.

Protocol: Above information was lacking in the protocol; to provide more information, we added the sentences above to the review.

Statistical analysis and data synthesis

Review: added the following: 'We wanted to differentiate between polyps and fibroids but a 2×2 table could not be prepared because several results were possible. Therefore, in the chosen analyses, we do not directly differentiate between polyps and fibroids but analyse them as polyp (vs no polyp) or submucous fibroid (vs no fibroid)'.

Analyses are the same in the protocol and in the review. In the review, we wanted to explain more clearly why we chose these analyses because we do not directly differentiate.

Review: added the following: 'Even with a low number of studies and/or sparse data (due to 100% sensitivity or specificity), the advice is still to use a hierarchical model (Takwoingi 2015). In case of fewer than five studies or if the models did not converge, we planned to use a univariate random-effects logistic regression model'.

We added information to the review to inform the reader about our decision to use a hierarchical model.

Review: added the following: 'For 3D SIS, we noted minimal variation in specificity and quite some variation in sensitivity. The models converged only if we assumed no correlation between logit sensitivity and logit specificity. This probably occurred because very little heterogeneity was evident, with almost all studies reporting specificity of 100%. If sensitivity varies and specificity does not, the assumption that there is no correlation may be valid'.

As advice indicates that a hierarchical model should be used, we had to assume no correlation between logit sensitivity and logit specificity; we added this information to the review.

Protocol: Results will be presented as summary sensitivity and specificity, and additionally as likelihood ratios.

Review: Likelihood ratios were not reported.

As we had several objectives, and as results were well presented with summary sensitivity and specificity, we decided that presenting results in likelihood ratios as well would not provide additional value and would have resulted in too much information to report.

Investigations of heterogeneity

Review: added the following: 'prior testing (prior testing or not) and whether evaluation of 2D or 3D SIS was blinded for clinical information'.

We preferred to additionally evaluate the effects of these variables and therefore have added them to the review.

As heterogeneity can also be assessed with forest plots and ROC plots, we added this information to the section.

Sensitivity analyses

Protocol: We will conduct sensitivity analyses to determine whether conclusions are robust to arbitrary decisions made regarding eligibility of studies and analyses performed.

Review: In this review, we explained how we performed planned sensitivity analyses: 'We evaluated what would have happened if we would have removed studies with a high risk of bias for either one of the QUADAS-2 domains'. Second, we performed a sensitivity analysis for the effect of the reference standard.

Protocol: Missing or uninterpretable data were classified as positive test results.

Review: In this review, we clarified 'data' and changed the sentence: In case of missing or uninterpretable index tests (failure of 2D SIS or 3D SIS), data were classified as positive test results.

Protocol: 'We also performed a sensitivity analysis for the effect of the reference standard'.

Review: We removed the sentence above because the reference standard is one of the QUADAS domains, and we (already) stated the following; 'We evaluated what would have happened if we would have removed studies with a high risk of bias for either one of the QUADAS-2 domains'. Studies with partial verification bias in the reference standard domain (not all women received the reference standard) were removed as part of sensitivity analyses to evaluate their influence on the main outcome.

INDEX TERMS

Medical Subject Headings (MeSH)

Endometrium [*diagnostic imaging]; Hysteroscopy [standards]; Leiomyoma [*diagnostic imaging]; Observational Studies as Topic; Polyps [*diagnostic imaging]; Prospective Studies; ROC Curve; Randomized Controlled Trials as Topic; Reference Standards; Sensitivity and Specificity; Sodium Chloride [*administration & dosage]; Solutions [administration & dosage]; Ultrasonography, Interventional [*methods] [standards]; Uterine Diseases [*diagnostic imaging]; Uterine Hemorrhage [etiology]

MeSH check words

Female; Humans