A systematic review and meta-analysis of the diagnostic accuracy of endometrial sampling tests for detecting endometrial cancer

Noha Sakna,[1] Marwa Elgendi,[1] Mohamed Hamed,[1] Ahmed Zeinhom,[1] Somia Labib,[1] Ashraf Nabhan[1]

[1] Department of Obstetrics and Gynecology, Faculty of medicine, Ain Shams University, Cairo, Egypt.

Abstract

Objectives: to determine the diagnostic accuracy of different endometrial sampling tests for detecting endometrial carcinoma.

Design: a systematic review and meta-analysis of studies of diagnostic accuracy.

Eligibility criteria: We included published cross-sectional studies of women, of all ages, who had an endometrial sampling for preoperatively detecting endometrial cancer with verification using histopathology of hysterectomy specimens as the reference standard. We excluded case control and case series studies.

Information sources: We searched the Cochrane library, MEDLINE/PubMed, CINAHL, Web of Science, and Scopus from the date of inception of the databases to December 31, 2022. We did not apply any restrictions on language or date of publication. We searched the references of included studies and other systematic reviews.

Risk of bias: We extracted study data and assessed study quality using the revised quality assessment tool for diagnostic accuracy studies (QUADAS-2).

Synthesis of results: We used bivariate diagnostic random-effects meta-analysis and presented the results in a summary receiver operating characteristic curve. We assessed the certainty of evidence as recommended by the GRADE approach.

Results: Ten included studies, published between 1986 and 2020, recruited 1390 participants. Seven studies were low risk of bias in all domains and all studies had low applicability concerns. The most commonly studies index tests were Pipelle and conventional dilation and curettage. The sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio (95% confidence intervals), for Pipelle were 0.774 (0.565, 0.900), 0.985 (0.927, 0.997), 97.000 (14.000, 349.000), and 0.241 (0.101, 0.442) and for conventional dilation and curettage were 0.880 (0.281, 0.993), 0.984 (0.956, 0.995), 59.300 (14.200, 153.000), and 0.194 (0.007, 0.732); respectively.

Conclusion: High certainty evidence indicates that pre-operative endometrial sampling particularly using Pipelle or conventional curettage is accurate in diagnosing endometrial cancer. Studies assessing other endometrial sampling tests were sparse.

Systematic review registration: Center for Open Science, osf.io/h8e9z

Keywords: Endometrial sampling; Endometrial carcinoma; Sensitivity; Specificity; Diagnostic accuracy

Article summary

Strengths and limitations of this study

- This systematic review included only high quality diagnostic test accuracy studies, specifically with low verification bias.
- This systematic review used QUADAS-2 tool for quality appraisal and the Grading of Recommendations, Assessment, Development and Evaluation recommendations for certainty of evidence.
- The meta-analysis provided a robust synthesis of the evidence of the diagnostic accuracy of endometrial sampling tests for women typically seen in clinical practice with suspected endometrial hyperplasia or carcinoma.
- The possibility of unpublished studies always exists. The potential for publication bias was not assessed due to the lack of validated methods for diagnostic test accuracy reviews.

Introduction

Endometrial cancer (EC) is the fifth most common malignant condition among women worldwide. Its incidence shows an increase in line with an increase in life expectancy. It is a major health concern being the 14th leading cause of death from cancer in women.[1–5]

Most cases affect postmenopausal women.[6] Most women present with abnormal uterine bleeding.[7] While 90 percent of women with endometrial cancer initially present with PMB, only 5 to 10 percent of women with PMB are diagnosed with endometrial cancer. Irregular pre- or peri-menopausal bleeding is yet another important presenting symptom.[7] Endometrial cancer is sometimes diagnosed in women without an abnormal uterine bleeding. This occurs during the investigation of a thick endometrial line found on an imaging technique performed for other reasons or incidentally in hysterectomy specimens performed for benign conditions.[4,8]

Survival from endometrial cancer depends on the stage at diagnosis. By enabling an early detection of atypical hyperplasia and early-stage endometrial cancer, an accurate early detection remains the cornerstone for improving outcomes.[4,9,10]

Objectives

The aim of this systematic review and meta-analysis is to determine the diagnostic accuracy of different endometrial sampling tests for detecting endometrial carcinoma.

Methods

We conducted this systematic review using methodological approaches prespecified in a review protocol[11] which was prospectively registered in the Center for Open Science.[12] We reported this review according to the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) guidelines.[13]

Eligibility criteria

We included published cross-sectional studies that have evaluated the diagnostic accuracy of different endometrial sampling methods, followed by verification of study participants with a reference test. Participants are women, of all ages, who had an endometrial sampling for a suspected endometrial hyperplasia or cancer. Index tests included different endometrial sampling techniques. Target conditions

included endometrial hyperplasia or carcinoma. We accepted histopathology of hysterectomy specimens as the reference standard for the diagnosis of the target conditions. We excluded studies of a case-control design that included women with known endometrial carcinoma to matched controls.[14]

Information sources

We developed a comprehensive search strategy to find published articles[15]. We searched the Cochrane library, MEDLINE/PubMed, CINAHL, Web of Science, and Scopus from the date of inception of the databases to December 31, 2022. We did not apply any restrictions on language or date of publication.

Searching other resources

We manually searched the reference lists in articles retrieved from electronic databases and relevant review articles.

Study selection

Two review authors (SL and NS) used a reference manager, to import and de-duplicate search results. They independently screened titles and abstracts to identify potentially eligible studies for full-text retrieval. Two review authors (SL and MA) double checked the list of potentially eligible studies. Review authors working in pairs, independently assessed the full text using the predefined eligibility criteria. They resolved discrepancies by discussion and by consulting with the lead author (AN). We illustrated the study selection process using a PRISMA flowchart.

Data collection process

For each included study, two review authors (SL, MA), independently extracted the following data: General information: title, journal, year, and study design; Sample size: total number of participants included and tested; Baseline characteristics: age, index test, reference test; numbers of true positive, true negative, false positive, and false negative findings. We extracted these data for both target conditions (AEH and EC). All numbers were double checked by two authors (NS, AZ). We summarized the data from each study in 2 × 2 tables (true positive, false positive, false negative, true negative), and we stored the data into a spreadsheet.

Risk of bias and applicability

Review authors, working in pairs, independently assessed the risk of bias of included studies and applicability of their results using QUADAS-2.[16] They resolved discrepancies by discussion and by consulting with the lead author. We addressed aspects of study quality involving the participant spectrum, index tests, target conditions, reference standards, and flow and timing.[16] We classified a study as having a high risk of bias if at least one of the domains of QUADAS-2 was judged as being high risk.

Synthesis of results

Each method of endometrial sampling was compared against the reference test. For each method of endometrial sampling, TP, TN, FP, and FN were computed. We used bivariate diagnostic random-effects meta-analysis[17]. We presented the results in a summary receiver operating characteristic (ROC) curve. We used the Markov chain Monte Carlo procedure to generate summary positive and negative likelihood ratio and diagnostic odds ratio for the bivariate model[18]. When there are three or less studies, we used a univariate random-effects model. For each of the target condition, we presented the synthesis by the method of sample collection. We assessed the effect of risk of bias of included studies on diagnostic accuracy by performing a sensitivity analysis in which we excluded studies classified as having high risk of bias. We used R software version 4.2.2[19] and package 'mada'[20].

Summary of findings table and assessment of the certainty of evidence

We prepared summary of findings tables, using GRADEpro GDT[21], to present the main results and key information regarding the certainty of evidence. We assessed the certainty of evidence as recommended by the GRADE approach.[22]

We rated the certainty of evidence as either high (when not downgraded), moderate (when downgraded by one level), low (when downgraded by two levels), or very low (when downgraded by more than two levels) based on five domains: risk of bias, indirectness, inconsistency, imprecision, and publication bias. For each outcome, the certainty of evidence started as high when there were high-quality observational studies (cross-sectional or cohort studies) that enrolled participants with diagnostic uncertainty. If we found a reason for downgrading, we used our judgement to classify the reason as either serious (downgraded by one level) or very serious (downgraded by two levels)[23,24].

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Results

Study selection

Bibliographic database search identified 3,386 records and 23 additional records from searching the reference lists of relevant records. After removing duplicates, 2679 titles and abstracts were screened. Full text reports of 38 potentially eligible studies were assessed using the predefined eligibility criteria and ten studies were included, Figure 1.

Study characteristics

The eleven included studies, published between 1986 and 2020, had a cross-sectional design and recruited 1390 participants, Table 1.

The majority of recruited women were postmenopausal. Participants had different presentations including post menopausal bleeding, pre-menopausal abnormal uterine bleeding, or women scheduled for hysterectomy for different indications.

Studies used different methods of endometrial sampling. This included Pipelle[25–27], GDP Tao[28], Abradul[29], Li brush[30], Uterine Explora Curette (UEC)[31], Gynoscann[31], Novak[32], M curette[33] , or D&C[25,26,34]. The search did not identify an eligible study, as per the presepcified criteria for this review, that used an outpatient hysteroscopy-directed endometrial sampling. All studies used histopathology examination of hysterectomy as the reference standard. All studies reported EC and AEH.

Risk of bias and concerns regarding applicability

Regarding bias, eight studies were classified as low risk in all domains. One study was judged as high risk of bias in the domain of flow and timing. Two studies had an unclear risk of bias in the domains of patient selection. Three studies had an unclear risk of bias in the reference test domain. All studies had low concerns regarding applicability, Table 2.

Data synthesis

Diagnostic accuracy of endometrial sampling tests compared to histopathology of hysterectomy

Ten studies reported data on endometrial cancer and atypical endometrial hyperplasia using different methods of endometrial sampling, Table 3, Figure 2, and Figure 3.[25–34]

For endometrial cancer, out of 100 people with endometrial cancer, 80 would be correctly diagnosed and 20 would be incorrectly diagnosed. Out of 100 people without endometrial cancer, 99 would be correctly diagnosed and 1 would be incorrectly diagnosed, Table 3.

For atypical hyperplasia, out of 100 people with atypical hyperplsia, 76 would be correctly diagnosed and 24 would be incorrectly diagnosed. Out of 100 people without atypical hyperplasia, 98 would be correctly diagnosed and 2 would be incorrectly diagnosed, Table 3.

Subgroup analysis by the specific test

Three studies examined the accuracy of D&C for diagnosing EC.[25,26,34]. Using D&C, compared to histopathology examination of hysterectomy, out of 100 people with endometrial cancer, 88 would be correctly and out of 100 women without endomatrial cancer 98 would be correctly diagnosed, Table 4.

Three studies examined the accuracy of Pipelle for diagnosing AEH or EC.[25–27] Using Pipelle endometrial sampling, compared to histopathology examination of hysterectomy, out of 100 people with endometrial cancer, 77would be correctly diagnosed and out of 100 women without endomatrial cancer 99 would be correctly diagnosed, Table 4.

Diagnostic accuracy was examined in a single study for each of the following tests: Novak[32], GDP Tao device[28], Abradul cell sampler[29], Li brush[30], Uterine Explora Curette[31], Gynoscann[31], and Masterson Curette[33]. The measures ofdiagnostic accuracy are shown in Table 4.

Discussion

Summary of the evidence

This systematic review assessed the diagnostic accuracy of various endometrial sampling tests for detecting endoemtrial cancer or atypical endometrial hyperplasia. Only cross-sectional studies were considered for inclusion in the review. Ten DTA studies were identified, including 1390 women. All studies were published in peer-reviewed journals and most of these studies were of high methodological quality according to the QUADAS-2 criteria. Pooled estimates of diagnostic test accuracy were computed for various endometrial sampling tests. Robust evidence based on the results of the present review indicate that using Pipelle or conventional curettage can correctly confirm or exclude endometrial cancer or atypical endometrial hyperplasia. High quality evidence is insufficient for hysteroscopy, GDP Tao device, Abradul cell sampler, Li brush, Uterine Explora Cutette, Gynoscann, Z-sampler, Novac, and Masterson curette.

This systematic review represents the synthesis of exclusively high quality diagnostic accuracy studies. Cross-sectional studies in patients with diagnostic uncertainty and direct comparison of endometrial sampling test results with the appropriate reference standard (histopathology of hysterectomy) are considered high quality.

Every effort was made to minimize the risk of bias. First, we did not include case control studies or studies that include only women with endometrial carcinoma. Previously published reviews have included studies with case control design (two-gate design) which is generally not representative of

a test's accuracy in clinical practice.[9,35] Spectrum bias through case-control design would lead to suboptimal patient selection. This falsely inflates sensitivity and specificity. Further, both prevalence and predictive values depend on the ratio of people with and without the disease. In case–control studies, this ratio is constructed artificially, and thus prevalence and predictive values calculated from such a study are artefacts.[36–38,39,40] Other reviews have primary included studies of index tests in women with known endometrial carcinomna. The ability of the test to rule out the disease can never be calculated if the study recruited only participants known to have the condition.[40,41]

Second, we only included studies that used histopathology examination of hysterctomy as a reference test. Previous reviews has used multiple verification tests and tests that might not correctly classify the disease.[9,42,43] Errors in the reference test cause misclassification bias. Misclassification can significantly underestimate sensitivity and specificity. The magnitude of the bias depends on the disease prevalence, the accuracy of the index test, and the degree of misclassification. The misclassification rate can vary from study to study depending on the methodology associated with the reference test eg, skill of the pathologist.[39,44]

Limitations

A comprehensive literature search was conducted and a meticulous screening process was performed. In order to reduce the risk of publication bias, we did not implement any language or publication status restriction. However, it was not possible to search gray literature. The possibility of unpublished studies always exists. The potential for publication bias was not assessed due to the lack of validated methods for diagnostic test accuracy reviews.

Conclusions

In conclusion, pre-operative endometrial sampling, particularly using Pipelle or conventional curettage, is accurate in diagnosing endometrial cancer. There is insufficient high quality evidence regarding the diagnostic accuracy of other endometrial sampling tests.

Declarations

Funding statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Acknowledgments

We acknowledge the effort and the spirit of goodwill of Farida ElShafeey, Menna Kamel, and Aya Ramadan for their contribution to title and abstract screening.

Author contributions

AN conceived the idea for this review and designed the review methods.

NS, ME, SL collaborated in searching, screening, selecting studies, data extraction and synthesis.

AN wrote the first draft of the manuscript and NS, ME, AZ, MS, AN reviewed the manuscript.

All authors read and approved the final version of the manuscript.

Availability of data and materials

All data generated or analyzed during this study are included in the published scoping review article and are available in the project public repository.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

None declared.

References

1. Fitzmaurice C, Abate D, Abbasi N, Abbastabar H, Abd-Allah F, Abdel-Rahman O, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017. JAMA Oncology [Internet]. 2019 Dec 1;5(12):1749. Available from: http://dx.doi.org/10.1001/jamaoncol.2019.2996

2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians [Internet]. 2018 Nov;68(6):394–424. Available from: https://doi.org/10.3322/caac.21492

3. Concin N, Creutzberg CL, Vergote I, Cibula D, Mirza MR, Marnitz S, et al. ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. Virchows Archiv : an international journal of pathology. 2021 Feb;478:153–90.

4. Koskas M, Amant F, Mirza MR, Creutzberg CL. Cancer of the corpus uteri: 2021 update. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 2021 Oct;155 Suppl 1:45–60.

5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA: A Cancer Journal for Clinicians [Internet]. 2019 Jan;69(1):7–34. Available from: http://dx.doi.org/10.3322/caac.21551

6. Manchanda R, Thapa S. An overview of the main intrauterine pathologies in the postmenopausal period. Climacteric : the journal of the International Menopause Society. 2020 Aug;23:384–7.

7. Carugno J. Clinical management of vaginal bleeding in postmenopausal women. Climacteric : the journal of the International Menopause Society. 2020 Aug;23:343–9.

8. ACOG committee opinion no. 734: The role of transvaginal ultrasonography in evaluating the endometrium of women with postmenopausal bleeding. Obstetrics and gynecology. 2018 May;131:e124–9.

9. Hanegem N van, Prins MMC, Bongers MY, Opmeer BC, Sahota DS, Mol BWJ, et al. The accuracy of endometrial sampling in women with postmenopausal bleeding: A systematic review and meta-analysis. European journal of obstetrics, gynecology, and reproductive biology. 2016 Feb;197:147–55.

10. Visser NCM, Breijer MC, Herman MC, Bekkers RLM, Veersema S, Opmeer BC, et al. Factors attributing to the failure of endometrial sampling in women with postmenopausal bleeding. Acta obstetricia et gynecologica Scandinavica. 2013 Oct;92:1216–22.

11. Labib S, kamel M, Nabhan AF. Endometrial sampling tests for the diagnosis of endometrial carcinoma: A protocol of a diagnostic test accuracy systematic review [Internet]. 2022. Available from: http://dx.doi.org/10.31219/osf.io/vcwnt

12. Nabhan A. Diagnostic accuracy of different endometrial sampling tests for detecting endometrial carcinoma and its precursors. Open Science Framework [Internet]. 2022 Aug 30; Available from: https://osf.io/h8e9z/

13. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Group the PD. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA [Internet]. 2018 Jan;319(4):388–96. Available from: https://doi.org/10.1001/jama.2017.19163

14. Bossuyt P, Deeks J, Leeflang M, et al. Cochrane handbook for systematic reviews of diagnostic test accuracy version 2. London: Cochrane; 2022.

15. Nabhan A, Sakna NA, Elmawla MA, Labib S. searchRxiv [Internet]. 2023;2023. Available from: https://www.cabidigitallibrary.org/doi/abs/10.1079/searchRxiv.2023.00100

16. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011 Oct;155(8):529–36.

17. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. Journal of clinical epidemiology. 2005 Oct;58(10):982–90.

18. Zwinderman AH, Bossuyt PM. We should not pool diagnostic likelihood ratios in systematic reviews. Statistics in Medicine [Internet]. 2008 Feb;27(5):687–97. Available from: https://doi.org/10.1002/sim.2992

19. R Core Team. R: A language and environment for statistical computing. 2022; Available from: https://www.R-project.org/

20. Sousa-Pinto PD with contributions from B. Mada: Meta-analysis of diagnostic accuracy. 2022; Available from: https://CRAN.R-project.org/package=mada

21. GRADEpro G. Computer program. McMaster University (developed by Evidence Prime) GRADEpro GDT Hamilton (ON): McMaster University (developed by Evidence Prime). 2022;

22. Schünemann H, Brożek J, Guyatt G, Oxman editors A, et al. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October. 2013;2013.

23. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. Journal of clinical epidemiology. 2020 Jun;122:129–41.

24. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: Inconsistency, imprecision, publication bias, and other domains

for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. Journal of clinical epidemiology. 2020 Jun;122:142–52.

25. F B, E A. Diagnostic values of pipelle and standard curettage compared to hysterectomy pathology in postmenopausal bleeding: A comparative study. Advanced biomedical research [Internet]. 2020;9:58. Available from: https://pubmed.ncbi.nlm.nih.gov/33457341/

26. K G, O A, IE E, IA O, MM U, G Y, et al. Comparison of the histopathological diagnoses of preoperative dilatation and curettage and pipelle biopsy. European journal of gynaecological oncology [Internet]. 2014;35(5):539–43. Available from: https://pubmed.ncbi.nlm.nih.gov/25423700/

27. CR I, GS J, NK A. Study of the efficacy of pipelle biopsy technique to diagnose endometrial diseases in abnormal uterine bleeding. Journal of mid-life health [Internet]. 2019 Apr;10(2):75–80. Available from: https://pubmed.ncbi.nlm.nih.gov/31391756/

28. Leon MCB de, Wu HH, Lennon AE, Del Priore G. Novel approach to outpatient endometrial biopsy to detect endometrial cancer. The Journal of reproductive medicine [Internet]. 2016 May;61(5, 5-6):243–8. Available from: https://pubmed.ncbi.nlm.nih.gov/27424367/

29. G G, Velde J te. A new method for diagnostic curettage. European journal of obstetrics, gynecology, and reproductive biology [Internet]. 1987 Jul;25(3):239–42. Available from: https://pubmed.ncbi.nlm.nih.gov/3609437/

30. L H, J D, L Z, C S, Q W, X T, et al. An efficacious endometrial sampler for screening endometrial cancer. Frontiers in oncology [Internet]. 2019;9:67. Available from: https://pubmed.ncbi.nlm.nih.gov/30838173/

31. J K, I P, P SE, PE H, LG L, KL J, et al. Transvaginal ultrasound, endometrial cytology sampled by gynoscann and histology obtained by uterine explora curette compared to the histology of the uterine specimen. A prospective study in pre- and postmenopausal women undergoing elective hysterectomy. Acta obstetricia et gynecologica Scandinavica [Internet]. 1997 Sep;76(8):790–6. Available from: https://pubmed.ncbi.nlm.nih.gov/9348260/

32. Stever MR, Farmer G, Hernandez E, Miyazawa K. Routine prehysterectomy endometrial biopsy in a series of 523 women. The Journal of the American Osteopathic Association. 1986;86(9):67–9.

33. L R, E H, G B, D H, L A, PB H. The effectiveness of the masterson curette in sampling the endometrial cavity. Journal of the National Medical Association [Internet]. 1995 Dec;87(12):877–80. Available from: https://pubmed.ncbi.nlm.nih.gov/8558620/

34. A S, A M, A Z, S J, A Z. Diagnostic accuracy of endometrial curettage in endometrial pathology. Journal of Ayub Medical College, Abbottabad : JAMC [Internet]. 2011 Jan;23(1):129–31. Available from: https://pubmed.ncbi.nlm.nih.gov/22830167/

35. Dijkhuizen FPHLJ, Mol BWJ, Brolmann HAM, Heintz APM. The accuracy of endometrial sampling in the diagnosis of patients with endometrial carcinoma and hyperplasia. Cancer [Internet]. 2000 Oct 15;89(8):1765–72. Available from: http://dx.doi.org/10.1002/1097-0142(20001015)89:8<1765::aid-cncr17>3.0.co;2-f

36. Whiting PF, Rutjes AWS, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. Journal of Clinical Epidemiology [Internet]. 2013;66(10):1093–104. Available from: https://www.sciencedirect.com/science/article/pii/S089543561300200X

37. Leeflang MMG. Systematic reviews and meta-analyses of diagnostic test accuracy. Clinical Microbiology and Infection [Internet]. 2014;20(2):105–13. Available from: https://www.sciencedirect.com/science/article/pii/S1198743X14600434

38. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PMM, Kleijnen J. Sources of Variation and Bias in Studies of Diagnostic Accuracy. Annals of Internal Medicine [Internet]. 2004 Feb 3;140(3):189. Available from: http://dx.doi.org/10.7326/0003-4819-140-3-200402030-00010

39. Whiting PF, Rutjes AWS, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. Journal of Clinical Epidemiology [Internet]. 2013 Oct;66(10):1093–104. Available from: http://dx.doi.org/10.1016/j.jclinepi.2013.05.014

40. Hall MK, Kea B, Wang R. Recognising bias in studies of diagnostic tests part 1: Patient selection. Emergency Medicine Journal. 2019;36(7):431–4.

41. Willis BH. Spectrum bias–why clinicians need to be cautious when applying diagnostic test studies. Family Practice [Internet]. 2008 Oct 1;25(5):390–6. Available from: http://dx.doi.org/10.1093/fampra/cmn051

42. Visser NCM, Reijnen C, Massuger LFAG, Nagtegaal ID, Bulten J, Pijnenborg JMA. Accuracy of Endometrial Sampling in Endometrial Carcinoma. Obstetrics & Gynecology [Internet]. 2017 Oct;130(4):803–13. Available from: http://dx.doi.org/10.1097/AOG.0000000000002261

43. Narice BF, Delaney B, Dickson JM. Endometrial sampling in low-risk patients with abnormal uterine bleeding: a systematic review and meta-synthesis. BMC Family Practice [Internet]. 2018 Jul 30;19(1). Available from: http://dx.doi.org/10.1186/s12875-018-0817-3

44. Schmidt RL, Factor RE. Understanding Sources of Bias in Diagnostic Accuracy Studies. Archives of Pathology & Laboratory Medicine [Internet]. 2013 Apr 1;137(4):558–65. Available from: http://dx.doi.org/10.5858/arpa.2012-0198-RA

Table 1: Characteristics of included studies

Author	Year	Design	Outcomes	Sample size	Index test	Reference test	Participants
Behnamfar[25]	2020	Cross-sectional DTA	AEH, EC	25	Pipelle, D&C	Histopathology of hysterectomy	All women were menopausal, with post menopausal bleeding (mean age 60)
de Leon[28]	2016	Cross-sectional DTA	AEH, EC	42	GDP Tao	Histopathology of hysterectomy	Pre and post-menopausal women, mostly with abnormal uterine bleeding (median age 57)
Gerretsen[29]	1987	Cross-sectional DTA	AEH, EC	61	Abradul	Histopathology of hysterectomy	Pre and post-menopausal women who had a total hysterectomy
Gungorduk[26]	2014	Cross-sectional DTA	AEH, EC	189	Pipelle, D&C	Histopathology of hysterectomy	Pre and post-menopausal women who had a total hysterectomy (mean age 49)
Han[30]	2019	Cross-sectional DTA	AEH, EC	271	Li brush	Histopathology of hysterectomy	Pre and post-menopausal women who had a total hysterectomy
Ilavarasi[27]	2019	Cross-sectional DTA	AEH, EC	106	Pipelle	Histopathology of hysterectomy	Pre and post-menopausal women with AUB who had a total hysterectomy
Kufahl[31]	1997	Cross-sectional DTA	AEH, EC	181	Gynoscann, UEC	Histopathology of hysterectomy	Pre and post-menopausal women who had a total hysterectomy (median age 48)
Reddington[33]	1995	Cross-sectional DTA	AEH, EC	25	Masterson curette	Histopathology of hysterectomy	Pre and post-menopausal women who had a total hysterectomy
Saadia[34]	2011	Cross-sectional DTA	AEH, EC	50	D&C	Histopathology of hysterectomy	Pre and post-menopausal women who had a total hysterectomy (median age 43)
Stever[32]	1986	Cross-sectional DTA	AEH, EC	440	Novak	Histopathology of hysterectomy	Pre and post-menopausal women who had a total hysterectomy (mean age 40)

Table 2: Quality assessment of diagnostic accuracy studies (QUADAS-2)

Study ID	Risk of bias	Concerns regarding

					applicability		
	Patient SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Behnamfar, 2020	Unclear	Low	Unclear	High	Low	Low	Low
de Leon, 2016	Low	Low	Low	Low	Low	Low	Low
Gerretsen, 1987	Low	Low	Low	Low	Low	Low	Low
Gungorduk, 2014	Low	Low	Low	Low	Low	Low	Low
Han, 2019	Low	Low	Low	Low	Low	Low	Low
Ilavarasi, 2019	Low	Low	Low	Low	Low	Low	Low
Kufahl, 1997	Low	Low	Low	Low	Low	Low	Low
Reddington, 1995	Low	Low	Low	Low	Low	Low	Low
Saadia, 2011	Low	Low	Unclear	Low	Low	Low	Low
Stever, 1986	Low	Low	Low	Unclear	Low	Low	Low

Table 3: GRADE Summary of findings table: Diagnostic accuracy of all endometrial sampling tests compared to histopathology of hysterectomy

Index test	**Target condition**	**Number of studies (participants)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**Positive LR (95% CI)**	**Negative LR (95% CI)**	**Number of results per 1,000 patients tested (95% CI) Assumed Prevalence 10% for endometrial cancer (Typically seen in women with clinical suspicion of endometrial pathology)**				**Certainty of the evidence**
							True positives will receive appropriate treatment	False negatives will be misdiagnosed and not receive appropriate treatment	True negatives will not undergo inappropriate treatment or unnecessary further testing	False positives will undergo inappropriate treatment	
All tests	Endometrial cancer	10 (1390)	0.798 (0.673, 0.884)	0.989 (0.981, 0.994)	76.000 (41.000, 130.000)	0.209 (0.118, 0.331)	80 (67 to 88)	20 (12 to 33)	890 (883 to 895)	10 (5 to 17)	⊕⊕⊕⊕ High
	Atypical hyperplsia	10 (1390)	0.760 (0.639, 0.849)	0.978 (0.953, 0.989)	36.100 (16.400, 69.900)	0.250 (0.155, 0.368)	76 (64 to 85)	24 (15 to 36)	880 (858 to 890)	20 (10 to 42)	⊕⊕⊕⊕ High

Table 4: Test accuracy for individual endometrial sampling tests for diagnosing endometrial cancer and atypical endometrial hyperplasia

Target condition	Index test	Studies	Sensitivity	Specificity	Positive LR	Negative LR
EC	Pipelle	3	0.774 (0.565, 0.900)	0.985 (0.927, 0.997)	97.000 (14.000, 349.000)	0.241 (0.101, 0.442)
	D&C	3	0.880 (0.281, 0.993)	0.984 (0.956, 0.995)	59.300 (14.200, 153.000)	0.194 (0.007, 0.732)
	Novak	1	0.167 (0.010, 0.806)	0.990 (0.974, 0.996)	26.400 (0.812, 108.00)	0.766 (0.196, 1.00)
	GDP Tao device	1	0.886 0.694 0.964	0.977 0.815 0.998	39.000 2.508 606.541	0.116 0.036 0.374
	Abradul cell sampler	1	0.875 0.396 0.987	0.992 0.924 0.999	103.250 6.375 1672.264	0.126 0.009 1.685
	Li brush	1	0.892 0.778 0.951	0.998 0.979 1.000	396.118 24.813 6323.572	0.108 0.049 0.238
	Uterine Explora Curette	1	0.676 0.440 0.847	0.997 0.972 1.000	224.588 13.835 3645.886	0.325 0.163 0.645
	Gynoscann	1	0.676 0.440 0.847	0.997 0.972 1.000	224.588 13.835 3645.886	0.325 0.163 0.645
	Masterson curette	1	0.917 0.517 0.991	0.976 0.808 0.998	38.500 2.463 601.770	0.085 0.006 1.214
AEH	Pipelle	3	0.738 (0.451, 0.906)	0.981 (0.951, 0.993)	42.600 (12.200, 109.000)	0.285 (0.096 0.561)
	D&C	3	0.804 (0.584, 0.923)	0.970 (0.703, 0.998)	57.900 (2.900, 307.000)	0.221 (0.089, 0.425)
	Novak	1	0.250 (0.034 0.762)	0.989 (0.973, 0.995)	28.400 (2.480, 94.900)	0.714 (0.241, 0.979)
	GDP Tao device	1	0.625 (0.219 0.908)	0.988 (0.891 0.999)	50 (2.872 870.56)	0.38 (0.107 1.346)
	Abradul cell sampler	1	0.917 (0.517 0.991)	0.991 (0.921 0.999)	104.5 (6.547 1667.926)	0.084 (0.006 1.195)
	Li brush	1	0.75 (0.365 0.94)	0.998 (0.982 1)	400.5 (24.173 6635.53)	0.25 (0.063 1.001)
	Uterine Explora Curette	1	0.938 (0.598 0.993)	0.98 (0.947 0.993)	46.875 (16.363 134.279)	0.064 (0.004 0.934)
	Gynoscann	1	0.938 (0.598 0.993)	0.934 (0.887 0.962)	14.266 (7.935 25.649)	0.067 (0.005 0.98)
	Masterson curette	1	0.833 (0.31 0.982)	0.979 (0.828 0.998)	40 (2.459 650.64)	0.17 (0.014 2.139)

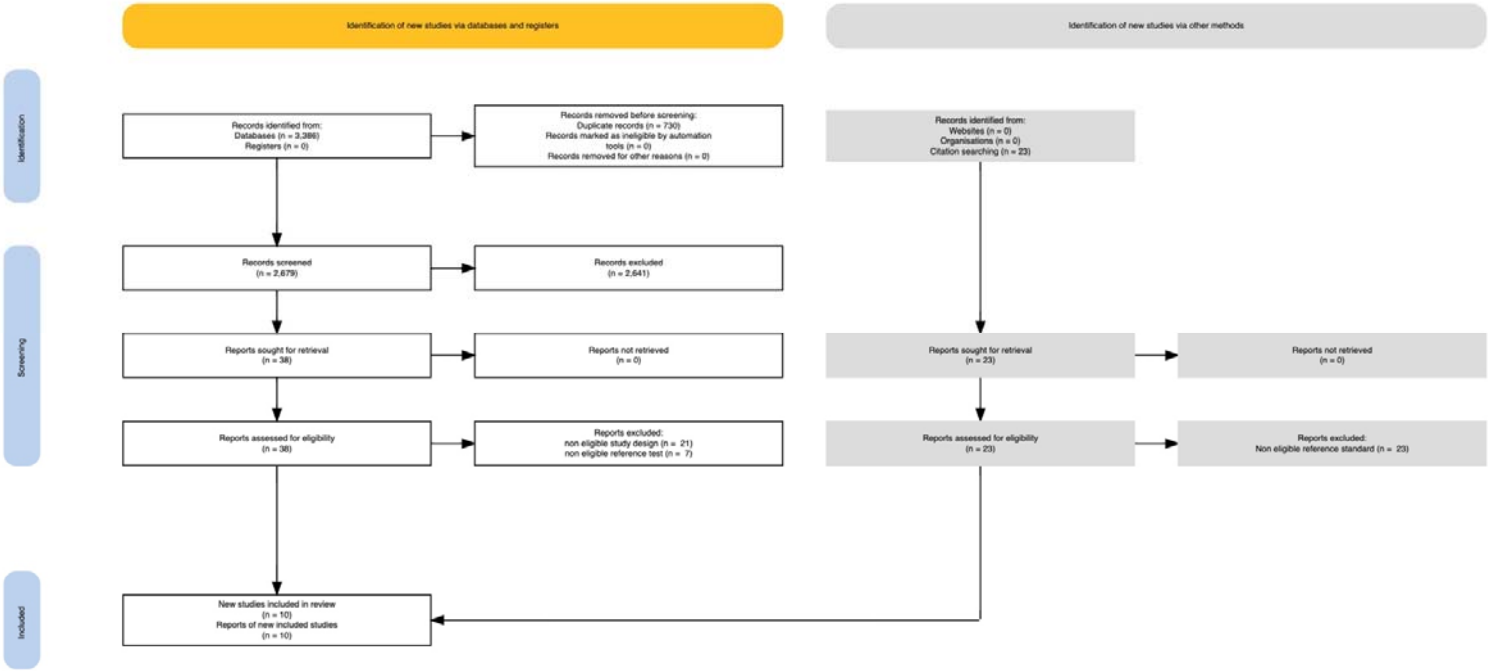

Figure 1 PRISMA Flow Diagram

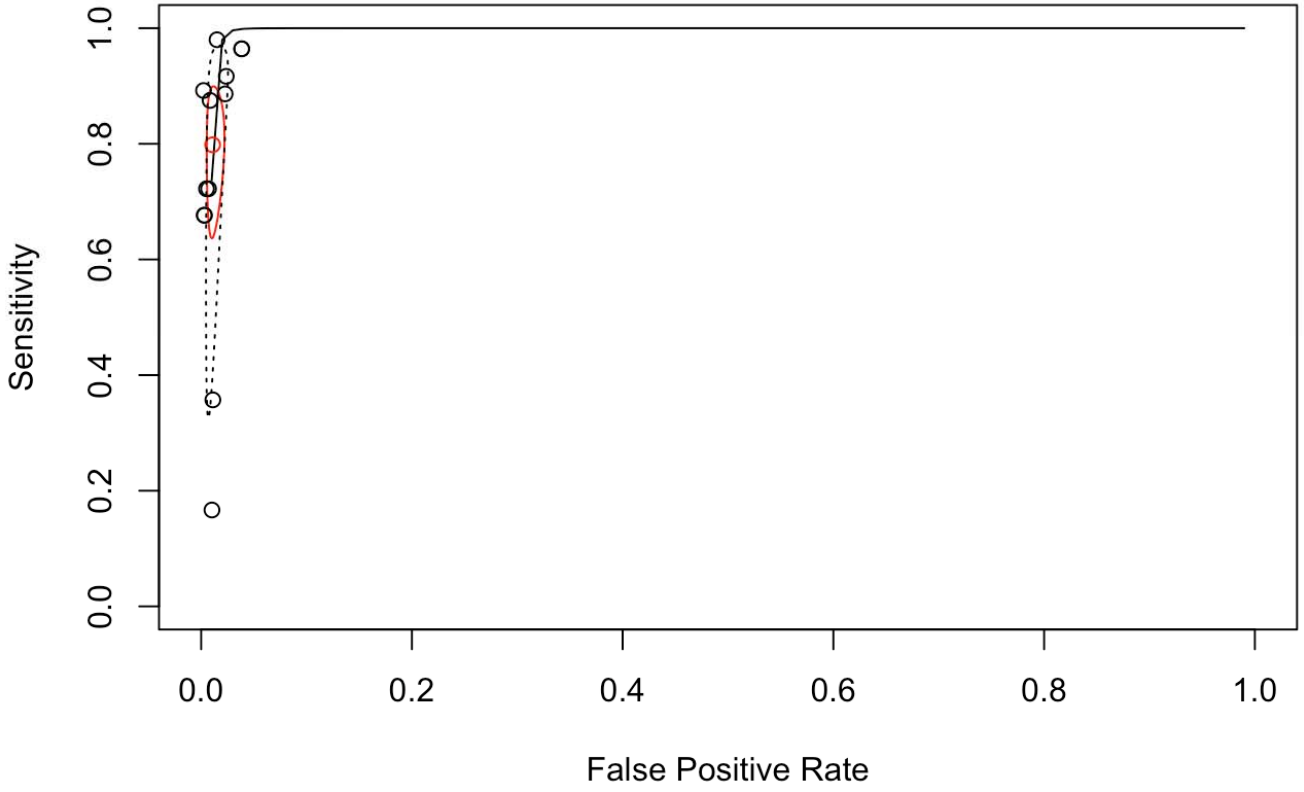

Figure 2 Summary receiver operating characteristic curve for all endometrial sampling tests for detecting endometrial carcinoma.

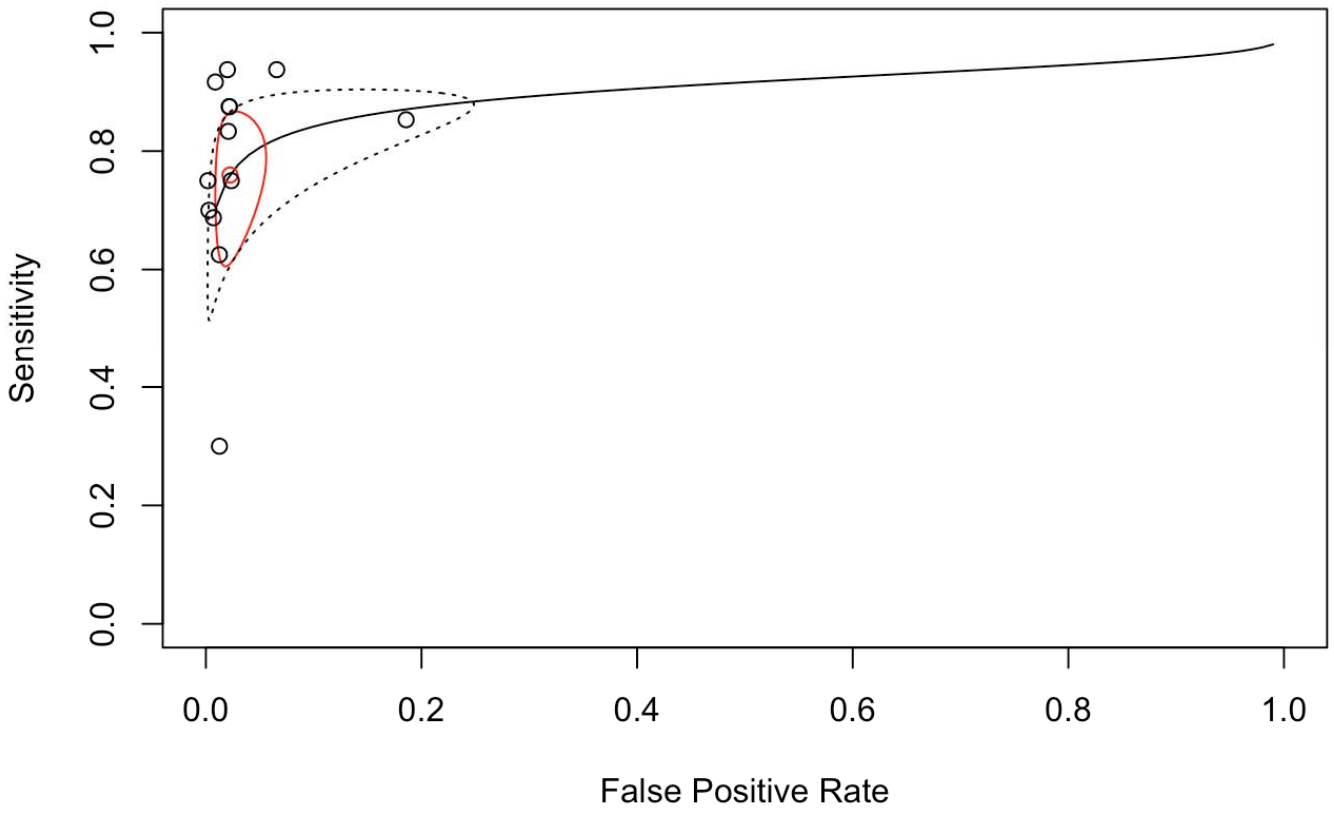

Figure 3 Summary receiver operating characteristic curve for all endometrial sampling tests for detecting atypical endometrial hyperplasia.