Endocervical sampling in women with suspected cervical neoplasia: a systematic review and meta-analysis of diagnostic test accuracy studies

Mathias Damkjær, MD; Jacob Brink Laursen, MD; Lone Kjeld Petersen, MedScD; Doris Schledermann, MD; Berit Bargum Booth, PhD; Britta Blume Dolleris, MD; Henrik Sehested Laursen, MLISc; Jeppe Bennekou Schroll, PhD

OBJECTIVE: Endocervical sampling in women with suspected cervical neoplasia can be performed by either endocervical brush or endocervical curettage. This study aimed to estimate the diagnostic accuracy, discomfort, and number of inadequate samples with either test.

DATA SOURCES: Four bibliographic databases were searched on June 9, 2022, with no date or language restrictions.

STUDY ELIGIBILITY CRITERIA: We included all diagnostic studies and randomized clinical trials that compared the endocervical brush with endocervical curettage in women with an indication for colposcopy.

METHODS: The review protocol was registered on the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42021222406). Two authors independently screened studies, extracted data, performed the risk-of-bias assessment (Quality Assessment of Diagnostic Accuracy Studies-2), and rated the certainty of the evidence using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) methodology. A meta-analysis of diagnostic test accuracy was performed using a bivariate random-effects model.

RESULTS: We included 7 studies: 4 diagnostic cohort studies and 3 randomized clinical trials. The reference standard was conization or hysterectomy. Risk of bias and concern about applicability were high for some of the studies in patient selection and flow and timing. Overall pooled sensitivity was 81% (95% confidence interval, 48—95; 799 women; 7 studies; low quality of evidence) for endocervical brush and 70% (95% confidence interval, 42—89; 761 women; 7 studies; low quality of evidence) for endocervical curettage. Overall pooled specificity was 73% (95% confidence interval, 36—93; 799 women; 7 studies; low quality of evidence) for endocervical brush and 81% (95% confidence interval, 56—94; 761 women; 7 studies; low quality of evidence) for endocervical curettage. The risk ratio for inadequate samples with endocervical curettage compared with endocervical brush was 2.53 (95% confidence interval, 0.58—11.0; $P= .215$; low-certainty evidence). Two studies reported on patient discomfort; one found less discomfort in the endocervical brush group, and the other found no difference.

CONCLUSION: No difference was found between endocervical brush and endocervical curettage in diagnostic accuracy, inadequate sampling rate, and adverse effects based on low-quality of evidence. Variation in the characteristics of women and the resulting diagnostic pathways make the external validity limited.

Key words: cervical cancer, diagnostic studies, diagnostic test accuracy meta-analysis, endocervical brush, endocervical curettage, endocervical neoplasia, endocervical sampling

Introduction

Cervical cancer is the fourth most frequent cancer in women⁴ and is a major cause of death worldwide, especially in low-income countries.⁵ It is well-established that cervical cancer is caused by infection with human papillomavirus (HPV) in epithelial cells, and the virus is found in approximately 99.7% of reported cervical cancers.⁶ HPV-induced cervical neoplasia is a precursor to cancer, gradually developing from mild cervical intraepithelial...
Different EB sampling devices exist, (single cells) from the endocervical canal, which provides a cytologic sample that follows the contours of the endocervical canal. This natural history of the disease has led to the implementation of primary prevention with HPV vaccines and secondary prevention with screening programs.

Cervical neoplasia is mostly asymptomatic, but women may present with postcoital bleeding. The strategy for diagnosing cervical neoplasia is most commonly colposcopy, biopsies from the transformation zone, and endocervical sampling. Indication for these diagnostic tests may be abnormal cytology, positive HPV DNA testing, or symptoms (e.g., postcoital bleeding). A positive finding (indicating high-grade neoplasia) on biopsies or endocervical sampling will result in further diagnostics, for example a colposcopy-guided conization. Two commonly used methods for obtaining an endocervical or cervical canal sample are the endocervical brush (EB) and endocervical curettage (ECC).

The EB is a collection device with a brush made of flexible plastic hairs that follow the contours of the endocervical canal, which provides a cytologic sample (single cells) from the endocervical surface. Different EB sampling devices exist, with a combined brush and spatula device being the most effective for collecting endocervical cells. ECC, on the other hand, is performed using a metal curette device that provides a histologic (tissue) sample. There is considerable variation in choice of test method and therefore it is relevant to conduct a review.

Objective
We aimed to systematically review the existing literature and perform a meta-analysis of diagnostic test accuracy (DTA) to provide estimates of the diagnostic accuracy of the EB and ECC, respectively, for the detection of cervical neoplasia in women with any indication for colposcopy, biopsies, and endocervical sampling. Furthermore, we investigated patient discomfort and inadequate sampling.

Materials and Methods

Eligibility criteria, information sources, search strategy
All diagnostic studies and randomized clinical trials (RCTs), regardless of setting, including women with abnormal cytology in cervical cancer screening and women with possible symptoms of cervical cancer undergoing colposcopy, biopsy, EB, or ECC sampling were included. The target condition was cervical neoplasia, and the reference standard was the final histologic result of either conization or hysterectomy. Furthermore, we also addressed the possible discomfort caused by either method.

We conducted a systematic search in the following 4 bibliographic databases on June 9, 2022, with no restrictions regarding language or date of publication: MEDLINE (Ovid), Embase (Ovid), the Cochrane Library, and CINAHL (EBSCO). The search strategy was originally developed by a medical librarian (H.S.L.) in collaboration with a subject advisor (L.K.P.) for a Danish national clinical guideline and was updated for the current review. To accommodate peer review, the original search strategy published in the protocol was modified to include studies without “colposkopy” as an obligatory search term, and >1000 more studies were screened with the search performed on June 9, 2022.

The modified search was a combination of “cervical” AND “neoplasia” AND “curettage,” with the corresponding synonyms in subject headings and title or abstract. The full search strategy can be found in Appendix 1. The search results were imported into Covidence for removal of duplicates and abstract screening.

Study selection
All studies were screened by 2 authors independently, and any disagreement was resolved through discussion.

Data synthesis
Two authors (B.B.B. and J.B.S.) extracted information about year, time period, author, country, funding, and threshold for treatment, and enough information was collected to create a 2-by-2 table (true positive, false positive, true negative, false negative).

Assessment of risk of bias
The same authors also individually performed bias assessment using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool and individually rated the certainty of evidence using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) methodology.

A full review protocol was published on February 24, 2021 on the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42021222406) before conducting this review. We adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.
## TABLE 1
Characteristics of the included studies in the quantitative analysis

<table>
<thead>
<tr>
<th>Study name</th>
<th>Country</th>
<th>Study design</th>
<th>Participants (number of women with histology present)</th>
<th>Inclusion criteria</th>
<th>Criteria for conization/ hysterectomy</th>
<th>Number of inadequate samples for diagnosis (curettage/ cytobrush)</th>
<th>Women discomfort</th>
<th>Definition of positive result</th>
<th>Funding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersen et al, 1988</td>
<td>United States</td>
<td>Cohort</td>
<td>87 (87)</td>
<td>All women who underwent conization in 1986 at that hospital</td>
<td>Inadequate colposcopy, positive ECC, suspension of microinvasion, earlier treatment failure, other</td>
<td>7/NR</td>
<td>NR</td>
<td>Endocervical dysplastic process</td>
<td>NR</td>
</tr>
<tr>
<td>Hoffman et al, 1993</td>
<td>United States</td>
<td>Cohort</td>
<td>388 (101 a)</td>
<td>Abnormal cytology</td>
<td>Inadequate colposcopy, CIN3, positive ECC, or cytobrush</td>
<td>11/2</td>
<td>NR</td>
<td>Endocervical involvement (single pathologist)</td>
<td>NR</td>
</tr>
<tr>
<td>Mogensen et al, 1997</td>
<td>Denmark</td>
<td>RCT</td>
<td>173 (88)</td>
<td>Abnormal cytology</td>
<td>≥CIN2</td>
<td>10/0</td>
<td>NR</td>
<td>≥CIN2</td>
<td>NR</td>
</tr>
<tr>
<td>Klam et al, 2000</td>
<td>Canada</td>
<td>RCT</td>
<td>315 (147)</td>
<td>Abnormal cytology</td>
<td>Biopsy-verified neoplasia (severity not specified)</td>
<td>0/1</td>
<td>No difference</td>
<td>Judgment of both tests by pathologist</td>
<td>NR</td>
</tr>
<tr>
<td>Boardman et al, 2003</td>
<td>United States</td>
<td>Cohort</td>
<td>62 (60)</td>
<td>Indication for conization or hysterectomy</td>
<td>Benign and (pre)malign</td>
<td>13/1</td>
<td>NR</td>
<td>Any degree of dysplasia, CIN, or malignancy</td>
<td>NIH (government)</td>
</tr>
<tr>
<td>Goksedef et al, 2013</td>
<td>Turkey</td>
<td>RCT</td>
<td>208 (33)</td>
<td>LSIL</td>
<td>Biopsy-verified neoplasia (severity not specified)</td>
<td>0/4</td>
<td>More pain with ECC</td>
<td>≥CIN1</td>
<td>NR</td>
</tr>
<tr>
<td>Frost, 1990</td>
<td>Denmark</td>
<td>Cohort</td>
<td>203</td>
<td>Abnormal cytology</td>
<td>≥CIN2 (or suspicion-based on clinical findings)</td>
<td>22/10</td>
<td>NR</td>
<td>≥CIN2</td>
<td>NR</td>
</tr>
</tbody>
</table>

CIN, cervical intraepithelial neoplasia; ECC, endocervical curettage; LSIL, low-grade intraepithelial lesion; NIH, National Institutes of Health; NR, not reported; RCT, randomized clinical trial.

a Only 101 included in the final analysis.

for Systematic Reviews and Meta-Analyses (PRISMA) for DTA\textsuperscript{15} and the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.\textsuperscript{13}

Data synthesis

Meta-analysis was conducted using metafor package\textsuperscript{14} and meta package\textsuperscript{15} in R, version 4.1.1 (R Foundation for Statistical Computing, Vienna, Austria). In accordance with the recommendations from the Cochrane DTA handbook, we conducted a meta-analysis if ≥3 studies without substantial clinical heterogeneity were available. We chose a bivariate random-effects model to determine summary estimates of sensitivity and specificity with 95% confidence intervals (CIs), and a univariate random-effects model to estimate risk ratios (RRs) for inadequate samples with 95% CI and P value with a significance level of α = 0.05.

A sensitivity analysis was performed to validate the credibility of the results, excluding studies at high risk of bias in QUADAS-2 domains, as described in our protocol. A separate analysis was done for each domain in turn. Furthermore, a separate analysis stratifying by chosen study design was performed.

Results

Study selection

We identified 2636 studies, and after removal of 1014 duplicates, 2 authors screened 1622 studies. We excluded 1506 studies after title and abstract screening and obtained 116 studies for full-text screening, 4 of which were from additional sources. Seven studies\textsuperscript{16–22} were ultimately included in the meta-analysis. The studies were conducted from 1988 to 2013, with a total of 1097 women included in the final analysis. The study characteristics are shown in Table 1. We excluded 109 studies for the following reasons: wrong study design (n=58), wrong intervention (n=27), wrong comparator (n=10), not available (n=6), wrong setting (n=4), ongoing trials (n=2), and wrong outcomes (n=2) (Figure 1).

Study characteristics

The 7 studies were performed in the United States, Denmark, Turkey, and Canada and included between 62 and 388 women. Three studies were RCTs\textsuperscript{18,20,21} and 4 were cohort studies,\textsuperscript{16,17,19} where the study by Boardman et al\textsuperscript{17} had the sequence of index tests randomized (table—envelope).

In the RCT by Mogensen et al,\textsuperscript{21} women undergoing colposcopy-directed biopsies were randomized to either EB or ECC. Women with ≥CIN2 underwent conization, and 131 of 180 women were included in the final analysis. Women with ≤CIN1 or an inadequate sample underwent follow-up with biopsies in a crossover design where previous EB changed to ECC and vice versa. The RCT by Klam et al\textsuperscript{20} randomized women to either EB or ECC, and both index tests were assessed by a single pathologist. Only 147 of 315 women with biopsy-verified disease underwent conization. In addition, in the RCT by Goksedef et al\textsuperscript{18} women were randomized to either EB or ECC, and both index tests were assessed by a single pathologist. However, in this study only 33 of 208 women with biopsy-verified disease underwent cone biopsy.

In all 4 cohort studies participants had both index tests done and were compared with either conization or hysterectomy. One of the cohort studies included women with conization performed at a single center,\textsuperscript{17} and the other included women with any indication for conization or hysterectomy.\textsuperscript{17} In 3 RCTs and 1 cohort study, women with an abnormal cytology were included.\textsuperscript{18–21} A considerable variation in the prevalence of cervical neoplasia was present in these studies, ranging from 6%\textsuperscript{22} to 90%.\textsuperscript{6}

The results of sensitivity and specificity for the individual included studies are shown in Figure 2. Detection of ≥CIN 2 was chosen as the cutoff value; however, some studies did use other cutoffs (Table 1). The highest sensitivity reported was 97% for the EB and 94% for ECC.\textsuperscript{22} The highest specificity was 97% for the EB and 97% for ECC.\textsuperscript{17}

Risk of bias of included studies

The QUADAS-2 scores are shown in Table 2. All studies were rated as having unclear risk of bias in patient selection because the index tests were assessed in 2 settings—one where women were undergoing conization (2 of the cohort studies\textsuperscript{16,17}) and one based on abnormal cytology referral (3 RCTs and 1 cohort study\textsuperscript{18–21}). However, in both settings only women who underwent conization (reference standard) were included in...
the final analysis. Four studies\textsuperscript{16–18,22} were rated to have concerns of applicability because of a more selected patient group than our predefined target population, which included all women with an indication for colposcopy, biopsies, and endocervical sampling (patient selection; QUADAS-2, domain 1). The index tests were generally assessed to have low risk of bias, but there were some applicability issues related to our research question. Klam et al and Goksedef et al\textsuperscript{18,20} were rated to have high risk of concerns regarding applicability because they used the same criteria for both ECC and EB (interpretation by a pathologist), which does not correspond to clinical practice (index test; QUADAS-2, domain 2). Most studies had insufficient details on the reference standard. Klam et al,\textsuperscript{20} Mogensen et al,\textsuperscript{21} and Goksedef et al\textsuperscript{18} were RCTs in a clinical setup, and therefore not all women had a conization or hysterectomy done. Klam et al\textsuperscript{20} was rated as having high risk of bias and applicability concerns regarding the reference standard because women without a reference test conducted were classified as true negatives. The rest of the included studies were rated as having unclear risk of bias and applicability concerns regarding the reference standard because different diagnostic criteria were used for conization and hysterectomy across the studies (QUADAS-2, domain 3). The criteria are shown in Table 1. The following 3 studies were rated as having high risk of bias in flow and timing. Firstly, in Mogensen et al\textsuperscript{21} only 88 of 173 women had conization done, and 49 women were excluded from analysis after randomization. Secondly, in Klam et al\textsuperscript{20} only 147 of 308 women underwent conization, but women without a reference standard were still included in the analysis and coded as true negatives. Thirdly, in Goksedef et al\textsuperscript{18} only 33 of 208 women underwent conization, and 40 of 248 women were excluded from the study (QUADAS-2, domain 4).

Synthesis of results
We performed a DTA meta-analysis for summary sensitivity and specificity for all 7 included studies. The forest plots for our main analysis are shown in Figures 2 and 3. The sensitivity analysis and the Bayesian receiver-operating characteristics curves are shown in the supplied Appendix 2 and 3.

The overall pooled sensitivity was 81% (95% CI, 48–95; 799 women; 7 studies) for EB and 70% (95% CI, 42–89; 761 women) for ECC.

The overall pooled specificity was 73% (95% CI, 36–93; 799 women) for EB and 81% (95% CI, 56–94; 761 women) for ECC.

The sensitivity analysis did not alter the interpretation of the results; however, in the analysis sorted by study design, the sensitivities of the 2 methods were equivalent. In the sensitivity analysis on risk of bias, no studies were excluded in QUADAS-2 domain 1 (all studies had unclear risk of bias) and in domain 2 (all studies
had low risk of bias). One study was omitted in QUADAS-2, domain 3 because it was found to have high risk of bias given that women without a reference standard were assigned to the true negative category. Three included studies were omitted in QUADAS-2, domain 4 because not all included women had the reference standard assigned. When sorted by study design, the 3 RCTs reported equivalent sensitivity for both EB and ECC (Appendix 2).

We also performed a meta-analysis of summary RRs for risk of inadequate samples. The overall RR was 2.53 (95% CI, 0.58–11.0; *P* = .215) for ECC compared with EB for inadequate samples (Figure 4).

Two studies reported on discomfort, and only the study by Goksedef et al reported a significant difference between ECC (mean [standard deviation] visual analog scale [VAS] score, 2.55 [1.12]) and EB (mean [standard deviation] VAS, 1.99 [0.87]).

We rated the evidence using the GRADE tool for diagnostic testing. We downgraded the quality of evidence from high to low: one level because of inconsistency in study designs because studies had different criteria for conducting either hysterectomy or conization, and another level because of indirectness because of differences in included study population and because of imprecision in the estimates with wide CIs. The main results and the gradings are shown in Tables 3 and 4.

### Comment

**Main findings**

We analyzed the results of 7 studies reporting the DTA of EB and ECC, and including data for 799 women undergoing EB and 761 women undergoing ECC.

The overall sensitivity was found to be slightly higher for EB than for ECC (low-certainty evidence), and vice versa, specificity was found to be slightly higher for ECC than for EB (low-certainty evidence). The results, however, were nonsignificant.

Furthermore, a nonsignificant increased risk of inadequate samples with ECC compared with EB was found, and no clinical difference in discomfort of the women between the tests were found.

As reported, EB was found to have a nonsignificant higher sensitivity, but lower specificity than ECC. It has been argued that because of this lower specificity EB is not as valuable as ECC as a diagnostic test, but is rather a superior screening tool.

If assuming a cervical neoplasia prevalence of 50% for women undergoing endocervical sampling, and using our pooled sensitivity and specificity estimates for EB and ECC, EB would detect 6 more cases of cervical neoplasia relative to ECC for every 100 women undergoing endocervical sampling (Tables 3 and 4). However, EB would also lead to 4 more false-positives compared with ECC. Although conization is considered safe, it is still associated with side effects such as increased perinatal mortality owing to preterm delivery, stenosis, and a very small risk of excessive bleeding.

The above estimates are based on previous testing and chosen populations in the included studies. The diagnostic accuracy may differ in a population with, for instance, only endocervical high-grade squamous intraepithelial lesion (HSIL) cytology and negative biopsies. A concern of too many false-positives (too low specificity) may limit the use of EB as a sole diagnostic test in this population for referral to a diagnostic conization. Accuracy in this situation must be addressed empirically because it was only in the study by Hoffman et al that endocervical HSIL in itself elicited a diagnostic conization. The diagnostic pathway for these women differs globally. In the Australian and Canadian guidelines, for women with negative colposcopy-guided biopsies but with HSIL-positive cytobrush samples (and HPV-DNA-positive samples in Australia), a diagnostic excisional...
procedure should be considered. However, in other countries, Denmark included, this is not the case, and isolated HSIL only warrants a follow-up examination.

It has been suggested that ECC may be more sensitive for disease in the endocervical canal because its sampling technique is more precise than EB, and the high sensitivity and lower specificity of EB could be related to ectocervical disease (contamination). Modified sleeve technique has been proposed to increase the specificity of the EB by guiding and shielding the EB into the endocervical canal, thereby minimizing ectocervical contamination. However, in the studies by Hoffman et al, Mogensen et al, and Klam et al looking at verified endocervical disease, EB was still reported to have higher sensitivity for endocervical disease than ECC. In a study by Zahn et al, the interobserver agreement in the interpretation of ECC by blinded pathologists was only moderate, which could be considered slightly concerning.

In 2 of the included studies, 1 and 7 inadequate samples, respectively, have been registered as true-negatives instead of being excluded in analysis as seen in the other studies. This could lead to a false lower sensitivity but higher specificity for ECC because of misclassification. Yet, because of few inadequate samples in the 2 studies and because the rest of the included 5 studies excluded inadequate samples, this is not likely to be the reason for the diverging sensitivities of EB and ECC. In light of these findings, a recent study by Mihaljevic et al has reported in a subanalysis that ECC samples with abundant material had higher sensitivity than those with scant material. We would have liked to have included this study, but some errors and inconsistencies were found in their $2 \times 2$ tables, and the corresponding author did not reply to our e-mail. Another study found but not yet published was an RCT by Gonzalez et al that was presented in a conference paper in 2019 and has not undergone peer review. They did not publish the $2 \times 2$ tables, and we have not been able to contact the authors by e-mail.

Only 2 studies reported women’s discomfort. One study did not find any difference, and the other found a small association with slightly higher VAS score for ECC, but VAS of 2.55 vs VAS of 1.99 is not considered clinically relevant. Our results are in accordance with the RCT from Undurraga et al, reporting no difference in women’s preference for the 2 procedures.

Strengths and limitations
The strengths of this review include a prespecified protocol, comprehensive bias assessment using the QUADAS-2 tool, and a thorough literature search conducted by an expert research librarian.

This review had some limitations, including a low number of studies and participants. DTA studies vary in study
design because no clear superior study design has been proposed. Therefore, the included studies varied in design between RCT and diagnostic cohort studies, which caused heterogeneity. Another limitation is potential selection bias because only women who underwent a conization and/or hysterectomy were included in the final analysis of all but 1 study. However, these diagnostic procedures are considered the gold standard and as such are necessary for proper diagnostic accuracy measurements. As mentioned, the diagnostic pathway for women varies globally, leading to variable prevalence of cervical neoplasia in the included studies. This variation needs to be addressed in future studies. Furthermore, some of the studies were >20 years old, and although no major technical breakthrough regarding endocervical samples has emerged, a liquid-based analysis of cytology is now, for instance, the most used type of analysis, and HPV-DNA status is now also used in conjunction to colposcopy referral.

**Comparison with existing literature**

Our results are in accordance with the European Guidelines for Quality Assurance in Cervical Cancer Screening, Second Edition, reporting higher sensitivity for EB than for ECC, although also nonsignificant, leading to a higher detection rate for EB.

A real-life evidence study supported the use of ECC, especially in women aged >50 years. The finding that ECC may be more useful in older women is also supported by Gage et al. This may be because of a higher risk of inadequate biopsies in elderly women because the transformation zone is retracted with age into the cervical canal. ECC then might be more useful because it gives a histologic evaluation of the endocervical canal. However, these studies did not compare ECC with EB. A study also examined combining EB and ECC for endocervical sampling, reporting no difference in the amount of material attained from using both sampling techniques. However, their finding was in direct contrast to the findings by the newer RCT by Undurraga et al. More high-quality studies are needed to further address this idea of combined sampling.

Our results are in accordance with the RCT by Undurraga et al, which reported a higher risk of inadequate samples with ECC than with EB (14.3% vs 2.0%; \( P = .002 \)), and we also reported a higher, albeit nonsignificant, risk of inadequate samples using ECC. The finding of an inadequate sample rate of 14.3% in the study by Undurraga et al may be considered high. A study by Gage et al with >13,000 ECC samples only reported an inadequate sample rate of 4%. However, no clear definition of inadequate sampling was given in the studies.

**Conclusion and implications**

No difference was found between EB and ECC in diagnostic accuracy, inadequate sampling rate, and adverse effects based on low-quality of evidence. Variation in the characteristics of women and the resulting diagnostic pathways make the external validity limited.
**TABLE 4**
Overview of GRADE profile for endocervical curettage

**Endocervical curettage**
Population: women where there was an indication for endocervical diagnostic procedure
Index test: endocervical curettage
Reference standard: conization specimen or hysterectomy
Pooled sensitivity: 0.70 (0.42–0.89); and specificity: 0.81 (0.56–0.94)

<table>
<thead>
<tr>
<th>Test result</th>
<th>Number of women if 100 were tested</th>
<th>Number of participants (studies)</th>
<th>Certainty of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>True positives (correctly detected)</td>
<td>35 (21–45)</td>
<td>761 (7)</td>
<td>Low (● ● ○○)</td>
</tr>
<tr>
<td>False positives (test positive, but not cervical neoplasia)</td>
<td>9 (3–22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>True negatives (correctly excluded)</td>
<td>41 (28–47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>False negatives (missed)</td>
<td>15 (5–29)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The prevalence of 50% was chosen as the mean prevalence of cervical neoplasia in the studies.

**GRADE**. Grading of Recommendations, Assessment, Development, and Evaluation.


**REFERENCES**
**Appendix 1**

**MEDLINE**

(Cervix Uteri/ or (cervix or cervical or vaginal or cervicovaginal or endocervi* or ectocervi*).ab,kf,ti.) and (exp Uterine Cervical Dysplasia/ or Cervical Intraepithelial Neoplasia/ or (dysplas* or neoplas*).ab,kf,ti.) and (exp Curettage/ or (curettage or ((cervi* or endocervi*) adj3 (abras* or abrad* or scrap*)) or cytobrush* or uterobrush* or endocervical sampler* or cervex-brush* or cyto-brush*).ab,kf,ti.)

**EMBASE**

((exp uterine cervix/ or (cervix or cervical or vaginal or cervicovaginal or endocervi* or ectocervi*).ab,kw,ti.) and (exp uterine cervix dysplasia/ or uterine cervix carcinoma in situ/ or (dysplas* or neoplas*).ab,kw,ti.) and (exp curettage/ or (curettage or ((cervi* or endocervi*) adj3 (abras* or abrad* or scrap*)) or cytobrush* or uterobrush* or endocervical sampler* or cervex-brush* or cyto-brush*).ab,kw,ti.) not (conference abstract or conference paper or "conference review" or editorial or letter).pt.

**CINAHL**

(((MH “Cervix”) OR (cervix OR cervical OR vaginal OR cervicovaginal OR endocervi OR ectocervi)) AND (((MH “Cervix Dysplasia”) OR ((MH “Cervical Intraepithelial Neoplasia”) OR (dysplas* OR neoplas*)) AND (((MH “Curettage+”) OR (curettage OR ((cervi* OR endocervi*) N3 (abras* OR abrad* OR scrap*))) OR cytobrush* OR uterobrush* OR endocervical sampler* OR cervex-brush* OR cyto-brush*)) OR (cervix-brush* OR endocervical sampler* OR cervex-brush* OR cyto-brush*))

---

**SUPPLEMENTAL TABLE 1**

<table>
<thead>
<tr>
<th>Search strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Search specialist</th>
<th>Henrik Laursen, Fagbibliotek, Regional Hospital Central Jutland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mail</td>
<td><a href="mailto:henlas@rm.dk">henlas@rm.dk</a></td>
</tr>
</tbody>
</table>

### SUPPLEMENTAL TABLE 2
#### Revised search strategy

<table>
<thead>
<tr>
<th>Databases</th>
<th>Interface</th>
<th>Hits</th>
<th>Date of search</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE</td>
<td>Ovid MEDLINE(R) ALL &lt;1946 to June 08, 2022&gt;</td>
<td>1099</td>
<td>Sept 6, 2022</td>
</tr>
<tr>
<td>Embase</td>
<td>OVID Embase 1974 to 2022 June 08</td>
<td>1179</td>
<td>Sept 6, 2022</td>
</tr>
<tr>
<td>CINAHL</td>
<td>EBSCO CINAHL with Full Text</td>
<td>274</td>
<td>Sept 6, 2022</td>
</tr>
</tbody>
</table>

Total number of references: 2634.

Number of references after deduplication in Covidence: 1620.


### SUPPLEMENTAL TABLE 3
#### COCHRANE Library

<table>
<thead>
<tr>
<th>#1</th>
<th>MeSH descriptor: [Cervix Uteri] explode all trees</th>
<th>1102</th>
</tr>
</thead>
<tbody>
<tr>
<td>#2</td>
<td>(cervix or cervical or vaginal or cervicovaginal or endocervi* or ectocervi*):ti,ab,kw</td>
<td>39,719</td>
</tr>
<tr>
<td>#3</td>
<td>#1 or #2</td>
<td>39,719</td>
</tr>
<tr>
<td>#4</td>
<td>MeSH descriptor: [Uterine Cervical Dysplasia] explode all trees</td>
<td>218</td>
</tr>
<tr>
<td>#5</td>
<td>MeSH descriptor: [Cervical Intraepithelial Neoplasia] explode all trees</td>
<td>609</td>
</tr>
<tr>
<td>#6</td>
<td>(dysplas* or neoplas*):ti,ab,kw</td>
<td>93,981</td>
</tr>
<tr>
<td>#7</td>
<td>#4 or #5 or #6</td>
<td>93,990</td>
</tr>
<tr>
<td>#8</td>
<td>MeSH descriptor: [Curettage] explode all trees</td>
<td>570</td>
</tr>
<tr>
<td>#9</td>
<td>(curettage or ((cervi* or endocervi*) near/3 (abras* or abrad* or scrap*)) or cytobrush* or uterobrush* or endocervical sampler* or cervex-brush* or cytobrush*):ti,ab,kw</td>
<td>1920</td>
</tr>
<tr>
<td>#10</td>
<td>#8 or #9</td>
<td>1920</td>
</tr>
<tr>
<td>#11</td>
<td>#3 and #7 and #10</td>
<td>82</td>
</tr>
</tbody>
</table>

- Cochrane reviews
- CENTRAL
- Cochrane Library search total

# Appendix 2

Sensitivity analysis

### SUPPLEMENTAL TABLE 4

Sensitivity analysis for endocervical brush and curettage

<table>
<thead>
<tr>
<th>Reason for exclusion</th>
<th>Number of studies analyzed</th>
<th>Endocervical brush</th>
<th>Endocervical curettage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pooled sensitivity in % (95% CI)</td>
<td>Pooled specificity in % (95% CI)</td>
</tr>
<tr>
<td>All studies</td>
<td>7</td>
<td>81 (48—95)</td>
<td>73 (36—93)</td>
</tr>
<tr>
<td>QUADAS-2 assessment</td>
<td>High risk in domain 1</td>
<td>7</td>
<td>Same result (all studies included)</td>
</tr>
<tr>
<td></td>
<td>High risk in domain 2</td>
<td>7</td>
<td>Same result (all studies included)</td>
</tr>
<tr>
<td></td>
<td>High risk in domain 3</td>
<td>6</td>
<td>79 (43—96)</td>
</tr>
<tr>
<td></td>
<td>High risk in domain 4</td>
<td>4</td>
<td>71 (30—94)</td>
</tr>
<tr>
<td>Study design</td>
<td>Randomized clinical trials</td>
<td>4</td>
<td>71 (30—94)</td>
</tr>
<tr>
<td></td>
<td>Cohort design</td>
<td>3</td>
<td>68 (25—97)</td>
</tr>
</tbody>
</table>

CI, confidence interval; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies-2.

Appendix 3
Bayesian SROC curve

**SUPPLEMENTAL FIGURE 1**
Bayesian SROC curve for endocervical brush

FPR, false positive rate; SROC, summary receiver operator characteristic; TPR, true positive rate.


**SUPPLEMENTAL FIGURE 2**
Bayesian SROC curve for endocervical curettage

FPR, false positive rate; SROC, summary receiver operator characteristic; TPR, true positive rate.