CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

HPV Self-Sampling for Primary Cervical Cancer Screening: A Review of Diagnostic Test Accuracy and Clinical Evidence

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	April 19, 2018
Report Length:	31 Pages

CADTH

Authors: Yi-Sheng Chao, Michelle Clark, Caitlyn Ford

Cite As: HPV Self-sampling for primary cervical cancer screening: A review of diagnostic test accuracy and clinical evidence. Ottawa: CADTH; 2018 Apr. (CADTH Rapid response report: summary with critical appraisal).

Acknowledgments:

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Abbreviations

AGC	atypical glandular cell
AIS	adenocarcinoma in situ
ASCUS	atypical squamous cells of undetermined significance, a common diagnostic threshold for positive cytology results [1]
CIN	cervical intraepithelial neoplasia, mostly diagnosed with colposcopy, regarded as precancerous changes, and further classified into CIN1, CIN2 and CIN3[1]
DOR	diagnostic odds ratio
DTA	diagnostic test accuracy
HC2	Hybrid Capture 2
HSIL	high-grade cervical lesions
LBC	liquid-based cytology
MALDITOF	matrix assisted laser desorption ionization-time of flight
NLR	negative likelihood ratio
HPV	human papillomavirus
NPV	negative predictive value
PCR	polymerase chain reaction
PLR	positive likelihood ratio
PPV	positive predictive value
RCT	randomized controlled trial
SR	systematic review

Context and Policy Issues

The importance of cervical cancer screening

In 2017, the incidence rates of cervical cancer in Canada ranged from 6.5 per 100,000 in Quebec to 10.7 per 100,000 in Newfoundland and Labrador.[2] In the past three decades, the age-standardized incidence rates have declined by 26%.[2] The reduction in cervical cancer incidence is mostly due to screening initiatives that aim to detect precancerous changes and cancer for treatment.[2] It is estimated that there has been an 80% reduction in the life-time risk of cervical cancer at the population level due to screening.[3]

The process of cervical cancer screening

Currently, the cervical cancer screening strategy in Canada is based on cytology and involves several steps. First, the cervical specimens are sampled and preserved.[2] Samples are applied to slides for cytological examination.[1] Liquid-based cytology is used when liquid media are used to preserve the samples and has the advantage of better diagnostic test accuracy.[4] Cytology or liquid-based cytology (LBC) is the primary method used in routine screening in Canada.[5] Positive cytological findings are defined by a threshold, especially atypical squamous cells of undetermined significance (ASCUS).[1]

People suspected to have positive results on the screening tests may be offered a second screening test or referred to colposcopy.[1] Colposcopy enables clinicians to directly inspect the cervix and a biopsy can be taken during the examination.[1] Based on the colposcopy findings, patients with precancerous or more advanced lesions, such as cervical intraepithelial neoplasia (CIN) 2 or a worse diagnosis, may be referred to treatment.[1]

HPV screening and the advantage of self-sampling

The human papillomavirus (HPV) tests directly detect HPV strains that are considered carcinogenic, especially HPV 16 and 18.[1] Depending on the types of HPV tests, HPV infection is detected based on various diagnostic thresholds.[1] There is a growing emphasis on the use of primary HPV tests in routine screening programs for the detection of precancerous changes in the cervix. HPV tests are more sensitive to precancerous changes and can be administered less frequently, every five years in several established screening programs,[1] compared to cytology every three years.[3] HPV tests can be conducted with self-collected or with clinician-collected samples.[1] There is evidence to show higher participation rates of self-sampled HPV tests than physician-sampled cytology.[6] Lastly, there are other potential advantages over cytology including reproducibility and cost-effectiveness.[7,8]

The importance of acceptance to reach the under-screened

A possible advantage of self-sampling is the potential for reaching underserved persons at risk of developing cervical cancer.[6] A 2013 meta-analysis showed higher participation rates of self-sampled HPV tests compared to physician-collected methods.[9] Currently the participation rate of physician-sampled cytology is not optimal in Canada according to the goals proposed by the Pan-Canadian Cervical Screening Network: no less than 80% of those eligible aged 21 years to 69 years should be screened in the preceding 42 months.[5] If corrected for hysterectomy, the participation rates range from 64.9% to 73.8% in British Columbia, Manitoba, and Ontario.[5]

Self-sampled HPV tests compared to physician-sampled tests

Self-sampled HPV tests require individuals to use brushes or other devices to collect samples from cervix by themselves.[1] The samples are used to determine HPV infection.[1] The adoption of self-collected samples raises the question, whether the gain in convenience is at the cost of diagnostic accuracy.[10] There is evidence showing that self-sampled HPV testing may have better sensitivity than cytology,[11] however, self-sampled HPV tests may not have equivalent diagnostic test accuracy (DTA) when compared with clinician-sampled HPV tests.[10] In two of the recently published systematic reviews, the DTA of clinician-sampled HPV tests has been evaluated against that of cytology.[1,4] However, the DTA of self-sampled HPV tests have not been well assessed against those of clinician-sampled tests.[4]

To understand the comparability and agreement of DTA between self- and clinician-sampled HPV tests, we aim to review the literature and compare the DTA between self- and clinician-sampled HPV tests or cytology.

Research Question

1. What is the diagnostic test accuracy of self-sampled HPV tests compared with clinician-sampled HPV tests or cytology for asymptomatic cervical cancer screening?
2. What is the clinical evidence regarding the agreement or concordance of self-sampled HPV tests and clinician-sampled HPV tests or cytology for asymptomatic cervical cancer screening?

Key Findings

There is evidence to show that self-sampled human papilloma virus (HPV) tests can achieve similar diagnostic test accuracy as clinician-sampled HPV tests with certain combinations of HPV tests and sampling devices for the detection of CIN2 (cervical intra-epithelial neoplasia) or severe diagnosis. For example, GP5+/6+ PCR HPV tests based on cervix specimens sampled with brushes or lavage have similar sensitivities and specificities as clinician-sampled HPV tests. Signal-based HPV tests including Hybrid Capture (HC2), one of the most widely tested HPV tests, are less sensitive and less specific with self-sampled specimens. There are individual studies showing high concordance or fair to high agreement between self- and clinician-sampled HPV tests. However, self-sampled HPV tests are less sensitive and specific than cytology at the threshold of ASCUS (atypical squamous cells of undetermined significance) or more severe dysplasia.

The advantages of self-sampled HPV tests included better acceptance by those eligible for routine screening programs. Self-sampled HPV tests detected more cases with findings of CIN2 or more severe than cytology or co-testing with clinician-sampled HPV tests and cytology.

The limitations of this review include considerable heterogeneity between studies, relatively few studies on the agreement between self- and clinician-sampled HPV tests, and the applicability of the existing evidence to vaccinated populations.

Methods

Literature Search Methods

A limited literature search was conducted on key resources including PubMed in Process, Medline via Ovid, Embase via Ovid, The Cochrane Library, University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and major international health technology agencies, as well as a focused Internet search. No filters were applied to limit the retrieval by study type. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2013 and March 14, 2018.

Selection Criteria and Methods

Two reviewers screened citations and selected studies in duplicate. The disagreement in inclusion or exclusion was resolved via meetings. In the first level of screening, titles and

abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Asymptomatic adults eligible for cervical cancer screening (≥ 21 years of age, or age at which screening starts in the jurisdiction)
Intervention	Q1-2: Self-sampled high-risk HPV tests for primary cervical cancer screening
Comparator	Q1-2: Clinician-sampled high-risk HPV tests for primary cervical cancer screening; cytology (conventional Pap smear or liquid based cytology) Q1 only: Colposcopy with histologic examination of tissue specimens, when indicated
Outcomes	Q1: Diagnostic test accuracy • Number and proportion of patients positive and negative on each test using colposcopy as reference standard • Sensitivity, specificity, PPV, NPV, PLR, NLR, DOR to screen for high-grade cervical lesions (HSIL or CIN2+, AGC, AIS) and/or invasive cervical cancer (squamous cell carcinoma or adenocarcinoma) Q2: Agreement between self-sampled HPV tests and clinician-sampled HPV tests or cytology (i.e., % agreement of positive test results, % agreement of negative test results)
Study Designs	Health technology assessments, systematic reviews/meta-analyses, randomized controlled trials, non-randomized studies

AGC = atypical glandular cell, AIS = adenocarcinoma in situ, CIN = cervical intraepithelial neoplasia, DOR = diagnostic odds ratio, HPV = human papillomavirus, HSIL = high-grade squamous intraepithelial lesion, NLR = negative likelihood ratio, NPV = negative predictive value, PLR = positive likelihood ratio, PPV = positive predictive value

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to 2013. Studies included in a selected systematic review were also excluded.

Critical Appraisal of Individual Studies

The included systematic reviews (SR) were critically appraised using the AMSTAR 2 tool.[12] Primary studies that investigated diagnostic test accuracy (DTA) of HPV tests were evaluated used the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) 2 instrument.[13] For the agreement between HPV tests, the quality of randomized clinical trials (RCTs) was assessed using the Cochrane Risk of Bias Tool.[14] The quality of non-randomized studies examining outcomes other than DTA was assessed using the Newcastle-Ottawa scale.[15] Summary scores were not calculated for the included studies; rather, a review of the strengths and limitations assessed in each included study were described.

Summary of Evidence

Quantity of Research Available

A total of 392 citations were identified in the literature search. Following screening of titles and abstracts, 353 citations were excluded and 39 potentially relevant reports from the electronic search were retrieved for full-text review. No potentially relevant publication was

retrieved from the grey literature search. Of these potentially relevant articles, 25 publications were excluded for various reasons, while 14 publications met the inclusion criteria and were included in this report. Appendix 1 presents the PRISMA flowchart of the study selection.

Summary of Study Characteristics

Additional details describing the characteristics of the included studies are reported in Appendix 2.

Study Design

One systematic review (SR) was identified for inclusion this report.[16] There were 34 unique studies included with data extracted from 36 articles used for meta-analysis.[16] One was a prospective cohort study, one retrospective cohort study, two RCTs, and 32 cross-sectional studies.[16] Sixteen included articles were screening studies, 17 follow-up studies, and three studies on high-risk populations.[16] One follow-up article was also using post-treatment samples.[16]

There were four RCTs and nine non-randomized studies.[16-29] Four non-randomized studies compared the diagnostic test accuracy (DTA) of self-sampled HPV tests with cytology.[19,22,25,26] All four non-randomized studies on the DTA were prospective cohort studies.[19,22,25,26] Two of the non-randomized studies not on the DTA were prospective cohort studies.[20,21] The other three non-randomized studies not on the DTA were cross-sectional studies.[23,24,29]

Year of Publication and Country of origin

The corresponding author of the SR published in 2014 was based in Belgium.[16] The four non-randomized studies on the DTA of self-sampled HPV tests published between 2014 and 2018were conducted in Denmark, the UK, Italy and the USA.[19,22,25,26] The RCTs were published between 2013 and 2018.[17,18,27,28] Two of the four RCTs were from Sweden and one each were from the USA and France.[17,18,27,28] The non-randomized studies not on the DTA were published between 2013 and 2017.[20,21,23,24,29] Two of the five non-randomized studies were from the USA and one each were from Denmark, the Netherlands, and France.[20,21,23,24,29]

Study population

The SR by Arbyn et al. did not have any restrictions on the literature search.[16] The 36 primary studies in Arbyn et al. included individuals eligible for cervical cancer screening programs, high-risk populations, and those being followed up after initial screening.[16] The primary studies in Arbyn et al. were conducted in low-, middle- and high-income countries.[16] The sample sizes ranged from 25 to 13,004.[16] Among 17 studies reporting mean ages, the mean ages ranged from 31 to 46.2 years.[16] No subpopulations were meta-analyzed.[16]

Among four non-randomized studies on the DTA of self-sampled HPV tests, the sample sizes ranged from 198 to 23,632.[19,22,25,26] The mean ages were 41.3 and 44.3 years in two studies in which it was reported.[22,25] Lam et al. focused on non-attendees, defined as individuals who were eligible but who had not undergone cervical cancer screening for at least four years.[19] The other three included people eligible for routine screening programs.[22,25,26]

The sample sizes of four RCTs ranged from 120 to 18,730.[17,18,27,28] The mean ages of all participants were not reported.[17,18,27,28] Gustavsson et al. and Sancho-Garnier et al. reported the results also by age groups.[17,28] Two of the four RCTs recruited non-attendees only.[27,28] Darlin et al. included those who were eligible for but who had not undergone screening for the past nine years.[27] Sancho-Garnier et al. enrolled people who had not responded to the first invitation for cytology screening.[28] In contrast, Gustavsson et al. conducted an RCT including people eligible for routine cervical cancer screening in Sweden.[17] Williams et al. included a population who had not been screened in the previous year.[18]

The sample sizes of the non-randomized studies not on the DTA ranged from 47 to 4801.[20,21,23,24,29] Lam et al. focused on the same "non-attendee" population as the other non-randomized study on the DTA by Lam et al.[20] Harvey et al. recruited those eligible for cervical cancer screening from temporary residential programs.[23] The other three studies enrolled people eligible for routine cervical cancer screening.[21,24,29]

Interventions and Comparators

The SR by Arbyn et al. compared the DTA between self-sampled HPV testing with clinician-sampled HPV testing for the detection of CIN2 or worse (CIN2+).[16] Arbyn et al. classified HPV tests into 13 types including Hybrid Capture 2 (HC2) (18 primary studies), PCRGP 5+/6+ (5), PCR-SPF10 (2), Aptima (1) and Cervista (1).[16] There were four types of sampling devices identified: brushes (18 studies), lavage (5), swabs (10), and tampons (1).[16] Colposcopy was considered the reference standard.[16]

The four non-randomized studies on the DTA compared self-sampled HPV testing with liquid-based cytology (LBC).[19,22,25,26] In addition, Stanczuk et al. also compared self-sampled HPV testing and LBC with clinician-sampled HPV testing.[22] Hybrid Capture 2 (HC2) was used in three studies.[19,25,26] Cobas was used in Stanczuk et al.[22] CLART and Onclarity were used in Lam et al.[19] Colposcopy was the reference standard.[19,22,25,26]

Two RCTs compared self-sampled HPV testing with cytology.[17,28] Another compared HPV self-sampling with clinician-sampled HPV testing[18] and the other with LBC.[27] RealTIme was used in Gustavsson et al. and Sancho-Garnier et al.[17,28] Cobas was used in Williams et al.[18] Luminex-based HPV test was used in Darlin et al.[27]

Four of the five non-randomized studies not on the DTA compared self-sampled HPV testing with clinician-sampled HPV testing[20,23,24,29] and the fifth compared it with cytology.[21] HC2 was used in Lam et al. and Harvey et al.[20,23] CLART was used in Lam et al.[20] Onclarity was used in Lam et al.[20] Cobas was used in Ketelaars et al.[21] INNO-LiPA HPV test was used in Haguenoer et al.[24] Linear Array was used in Castle et al.[29]

Outcomes

The primary outcomes reported in the SR by Arbyn et al. were the sensitivities and specificities of self- and clinician-sampled HPV testing for the detection of CIN2+ or CIN3+.[16]

Two of the non-randomized studies on the DTA of self-sampled HPV tests examined the sensitivities and specificities of self-sampled HPV testing for the detection of CIN2+ compared with cytology or LBC.[22,26]The other two reported the positive predictive values for the detection of CIN2+.[19,25]

Three RCTs compared the detection rates of CIN2+[17,27,28] and the other compared the agreement in HPV infection detection between self- and clinician-sampled HPV testing.[18]

Two non-randomized studies compared the detection rates of CIN2+[20,21] and the other three compared the agreement between self- and clinician-sampled HPV testing.[23,24,29]

Summary of Critical Appraisal

Additional details describing the critical appraisal of the included studies are reported in Appendix 3.

The SR by Arbyn et al. described the objectives and the rationale of the review.[16] Three databases that were searched to find relevant literature, PubMed, Embase, and CENTRAL.[16] Potential studies in all languages were considered.[16] There were no restrictions on the countries where the studies were conducted.[16] Independent duplicate study inclusion assessment and data extraction were performed.[16] The included primary studies were described.[16] The quality of included DTA studies were assessed with the QUADAS-2 checklist.[16] A bivariate model was applied to pool the sensitivities and specificities.[16] The heterogeneity due to countries, HPV sampling devices (brush, tampon, etc.) and HPV tests (HC2, Cobas, PCR-based methods, etc.) were tested.[16] There was a declaration of conflict of interest by the review authors.[16]

However, the review protocol was not published *a priori*.[16] Whether data was extracted in duplicate was unclear.[16] The excluded studies were not mentioned.[16] Funding sources of the primary studies not discussed.[16] The risk of bias of the included studies was not considered in meta-analysis.[21] The above-mentioned four weaknesses in four critical domains were identified and the confidence rating on the SR by Arbyn et al. was critically low.

Moreover, 16 of the 36 studies were based on screening populations, the sensitivities and specificities of screening populations were pooled with those of high-risk populations and those being followed up.[16] There was no sensitivity analysis to determine the impact of different populations on the pooled sensitivities and specificities.[16] This was different to the CADTH PICO criteria that focused on screening populations. The methodological quality and the populations used for meta-analysis were important limitations to draw conclusion for the research questions of this report.

Three of the non-randomized studies on the DTA described patient selection based on consecutive or random samples[19,22,25] and Jones et al. did not state the patient selection process.[26] The thresholds of the HPV tests were specified and self- or clinician-sampled HPV testing was conducted without the knowledge of the reference standard, colposcopy in these four studies.[19,22,25,26] The patients or clinicians were not blinded in the four studies as expected in the comparison between self- and clinician-sampled HPV tests.[19,22,25,26] The lengths of follow-up in the four studies were appropriate for DTA outcomes, ranging from 0.7 to 56.2 months.[19,22,25,26] Verification or ascertainment bias was adjusted in two studies.[22,26] Lam et al. only reported predictive values.[19]

Three of the four RCTs did not mention the method to randomly assign participants to different groups.[17,18,27] Sancho-Garnier et al. used software to generate a sequence for randomization.[28] ,There were no evidence regarding allocation concealment or blinding of patients and clinicians in the four RCTs.[17,18,27,28] However, the pathologists were blinded for group assignment in Gustavsson et al.[17] The lack of patient or clinician blinding were not considered a limitation for the comparison of self- and clinician-sampled HPV tests. Both

patients and clinicians were inevitably aware of the assignment. Patient attrition and loss to follow-up were reported in four RCTs.[17,18,27,28] The primary outcomes were reported in four RCTs.[17,18,27,28] Williams et al. adopted clinician-sampled HPV tests as reference standard,[18] while the other three used colposcopy.[17,27,28]

Four of the five non-randomized studies not on the DTA selected participants representative of the population in the target settings.[20,23,24,29] Ketelaars et al. did not describe how the cohort was derived.[21] Except for Lam et al. that used a cohort from another study as comparison,[20] the non-randomized studies conducted self-sampled HPV tests and physician-sampled tests on the same individuals, co-testing.[21,23,24,29] Without a clear description of the patient selection process, it was unclear whether individuals with known cervical lesions were used in two studies.[21,23] Lam et al. compared the two routine screening cohorts sampled in different time points both in the same region .[20] Each of the three co-testing studies used the same individuals to examine the effectiveness of intervention.[21,23,24,29]Every participant experienced both self- and clinician-sampled tests.[21,23,24,29] The lengths of follow-up were less than 16 months for the detection of HPV infection or CIN2+.[20,21,23,24,29]

Summary of Findings

1. *What is the diagnostic test accuracy of self-sampled HPV tests compared with clinician-sampled HPV tests or cytology for asymptomatic cervical cancer screening?*

Self-sampled HPV tests compared with clinician-sampled HPV tests

The SR by Arbyn et al. pooled the sensitivities and specificities found in 36 primary studies.[16] For the detection of CIN2+ or CIN3+, self-sampled HPV tests were less sensitive and specific than clinician-sample HPV tests based on the studies conducted in routine screening, high-risk, and follow-up populations in low- or high-income countries.[16] Signal-based assays, HC2 tested in 18 of the 36 included primary studies and Cervista in one study, were also found to be less sensitive and less specific with self-sampled specimens.[16] Aptima HPV tests with self-sampled specimens were less sensitive but not less specific than clinician-sampled tests.[16] By contrast, certain types of HPV tests (GP5+/6+ PCR, SPF10 PCR, Abbott Real Time high risk HPV test, DNA chip, modified GP5+/6+ PCR with Luminex reading, and matrix assisted laser desorption ionization-time of flight [MALDITOF]) were similarly sensitive and specific compared with self- and clinician-sampled specimens,.[16] The types of self-sampling devices seemed to be associated with differences in DTA.[16] Used with HC2 HPV tests, the self-sampling devices, brushes, swabs, and tampons, were associated with lower sensitivities than any self-sampling devices, while there was no significant difference observed for lavage.[16] For GP5+/6+ PCR, brushes and lavage were not associated with differences in sensitivities.[16] The DTA in the studies published in low- or middle-income was not significantly different from those in high-income countries.[16]

In Stanczuk et al., where authors adjusted for verification bias, Cobas HPV tests with self-sampled specimens were similarly sensitive and specific compared with clinician-sampled specimens.[22]

Self-sampled HPV tests compared with cytology

The SR by Arbyn et al. found that self-sampled HPV tests were less sensitive and specific than cytology at the threshold of atypical squamous cells of undetermined significance (ASCUS) for the detection of CIN2+[16] ASCUS was one of the cytological findings and often

considered a positive cytological result that required further management.[1] In the same SR, self-sampled HPV tests were similarly sensitive and less specific than cytology at the threshold of ASCUS+ for the detection of CIN3+.[16]

In Igidbashian et al., where authors did not adjust for verification bias, self-sampled HPV tests were found to be less specific than cytology in 708 samples.[25] Jones et al. adjusted for verification bias in 198 samples and suggested further testing of the self-sampling device in larger studies.[26]

2. *What is the clinical evidence regarding the agreement or concordance of self-sampled HPV tests and clinician-sampled HPV tests or cytology for asymptomatic cervical cancer screening?*

Self-sampled HPV tests were better accepted than clinician-sampled HPV tests in two RCTs.[27,28]

In addition to better acceptance, self-sampled HPV tests demonstrated high sensitivity and specificity using clinician-sampled HPV tests as reference standard in one RCT.[18] Two non-randomized studies also had the same findings.[23,24]

For the detection of HPV infection, a non-randomized study by Haguenoer et al. found that self-sampled HPV tests were as accurate in identifying the presence of HPV as clinician-sampled HPV tests.[24]

The results of self- and clinician-sampled Cobas HPV tests seemed concordant, 96.8% agreement, among the 2,049 participant samples in a prospective cohort study, Ketelaars et al.[21] Fair concordance, kappa = 0.54, was identified between tampon-based self-sampled and clinician-sampled HPV tests among 443 women .[29]

In terms of CIN2+ detection, self-sampled HPV tests showed higher detection rates than cytology that needed to be conducted by trained professionals and similar detection rates to HPV and cytology co-testing in one non-randomized study and one RCT.[19,28] In one RCT by Gustavsson et al., repeated self-sampled HPV tests were associated with two-fold higher detection rates of CIN2+ than cytology.[17]

The examination on the agreement between three types of HPV tests with self-sampled specimens, CLART, Onclarity and HC2, did not show good agreement, 27% agreement among three tests, and suggested further validation of HPV assays on self-sampled specimens.[20]

Limitations

There are several limitations to this report. First, there is considerable heterogeneity in the sampling devices, HPV assays, and clinical settings. In the SR by Arbyn et al., types of HPV assays and sampling devices were associated with different levels of DTA.[16] HC2 was the most extensively tested HPV assay in the SR and the primary studies.[16,19,20,23,25,26] However, there is evidence to show that the DTA of HC2 might be different from other HPV tests.[16,19] The sampling devices may also be associated with the differences in DTA.[16] This makes the direct comparisons between HPV tests difficult.

Second, the population used to draw conclusions in the SR by Arbyn et al. included not only screening populations, but also high-risk, follow-up and post-treatment populations.[16] Although a majority of the included studies were based on screening populations, 16 out of 36, there was no sensitivity analysis based on the types of populations.[16] Whether there

was a significant difference in DTA between screening populations and others remained unclear.

Lastly, there are fewer primary studies examining the agreement between self- and clinician-sampled HPV tests. Two of the studies showed conflicting results based on different HPV tests and populations.[21,29] The methods to define agreement also differed. For example, statistics like kappa,[26,29] sensitivities, specificities,[18,23,24] and concordance proportions[20,21] were used in different studies. This leads to some difficulty in comparing and integrating results.

Conclusions and Implications for Decision or Policy Making

Self-sampled HPV tests were similarly sensitive and specific to clinician-sampled HPV tests if certain types of HPV tests were used, such as Cobas (tested in one primary study[22]), GP5+/6+ PCR, and SPF10 PCR (both meta-analyzed in the SR by Arbyn et al.[16]). The most widely examined test, HC2, was less sensitive and less specific with self-sampled specimens.[16] For self-sampling devices used with HC2 HPV tests, brushes, swabs, and tampons were associated with lower sensitivities and lavage specimens were not.[16] For these devices used with GP5+/6+ PCR, brushes and lavage were not associated with differences in sensitivities.[16]

The SR by Arbyn et al. also suggested that self-sampled HPV tests were less sensitive and specific than cytology at the threshold of ASCUS+ for the detection of CIN2+.[16] Lower specificity than cytology was also found in another study.[25]

Self-sampled HPV tests were better accepted by those eligible for routine screening programs[27,28] and provided high sensitivities and specificities using clinician-sampled HPV tests as reference standard.[18,23,24] Only one study included underserved individuals.[18] Fair to high concordance between self- and clinician-sampled HPV tests was confirmed in two other studies.[21,29] Self-sampled HPV tests were as accurate as clinician-sampled tests to detect HPV infections.[24] However, the study examining three types of self-sampled versus clinician-sampled HPV tests did not show good agreement and might require further validation of self-sampled HPV tests.[20]

For the detection of CIN2+, self-sampled HPV tests identified more cases than cytology or co-testing with HPV tests and cytology.[17,19,28]

The limitations to this review included considerable heterogeneity between studies and relatively few studies examining the agreement between self- and clinician-sampled HPV tests.

For policy making, the applicability of the evidence remains unclear. This report focused on whether self- and clinician-sampled HPV tests were similarly accurate. Currently cytology remains the main screening strategy adopted in Canada and many countries. The transition from current practice to the adoption of self-sampled HPV tests requires consideration of context and population characteristics. The growing trend in HPV vaccination is further consideration when planning for new cervical cancer screening programs and policies.[1] The overall diagnostic accuracy of certain self-sampled HPV tests seems to be acceptable, however there is limited high quality evidence regarding the agreement between self-sampled and clinician-sampled testing. Further investigation is required.

References

1. Health technology assessment of human papillomavirus testing as the primary screening method for prevention of cervical cancer [Internet]. Mahon (IE): Health Information and Quality Authority; 2017. [cited 2018 Apr 5]. Available from: https://www.higa.ie/sites/default/files/2017-05/HPV%20HTA%20technical%20report-%2026052017_updated.pdf
2. Canadian cancer statistics 2017 special topic: pancreatic cancer [Internet]. Toronto (ON): Canadian Cancer Society; 2017. [cited 2018 Apr 5]. Available from: http://www.cancer.ca/~/media/cancer.ca/CW/publications/Canadian%20Cancer%20Statistics/Canadian-Cancer-Statistics-2017-EN.pdf
3. Cervical cancer - Patient algorithm [Internet]. Ottawa: Canadian Task Force on Preventive Health Care; 2018. [cited 2018 Apr 5]. Available from: http://canadiantaskforce.ca/wp-content/uploads/2016/05/2013-cervical-cancer-patient-algorithm-en.pdf
4. Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PP, Mustafa RA, et al. Cytology versus HPV testing for cervical cancer screening in the general population. Cochrane Database Syst Rev. 2017 Aug 10;8:CD008587.
5. Cervical cancer screening in Canada [Internet]. Toronto (ON): Canadian Partnership Against Cancer; 2016. [cited 2018 Apr 5]. Available from: https://content.cancerview.ca/download/cv/prevention_and_screening/cccic_microsite/documents/cccicmonitoringevalqualityindicatorspdf
6. Verdoodt F, Jentschke M, Hillemanns P, Racey CS, Snijders PJ, Arbyn M. Reaching women who do not participate in the regular cervical cancer screening programme by offering self-sampling kits: a systematic review and meta-analysis of randomised trials. Eur J Cancer. 2015 Nov;51(16):2375-85.
7. Introduction of molecular HPV testing as the primary technology in cervical cancer screening: Acting on evidence to change the current paradigm [Internet]. Toronto (ON): La Ki Shing Knowledge Institute, St. Michael's Hospital; 2015. [cited 2018 Apr 5]. Available from: http://healthydebate.ca/wp-content/uploads/2016/04/Report-on-HPV-primary-screening.pdf
8. El-Zein M, Richardson L, Franco EL. Cervical cancer screening of HPV vaccinated populations: Cytology, molecular testing, both or none. J Clin Virol [Internet]. 2016 Mar [cited 2018 Apr 5];76 Suppl 1:S62-S68. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4789074
9. Racey CS, Withrow DR, Gesink D. Self-collected HPV testing improves participation in cervical cancer screening: a systematic review and meta-analysis. Can J Public Health. 2013 Feb 11;104(2):e159-e166.
10. Arbyn M, Castle PE. Offering Self-Sampling Kits for HPV Testing to Reach Women Who Do Not Attend in the Regular Cervical Cancer Screening Program. Cancer Epidemiol Biomarkers Prev. 2015 May;24(5):769-72.
11. Zhao FH, Lewkowitz AK, Chen F, Lin MJ, Hu SY, Zhang X, et al. Pooled analysis of a self-sampling HPV DNA Test as a cervical cancer primary screening method. J Natl Cancer Inst [Internet]. 2012 Feb 8 [cited 2018 Apr 5];104(3):178-88. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3274511
12. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ [Internet]. 2017;358:j4008. Available from: http://www.bmj.com/content/bmj/358/bmj.j4008.full.pdf
13. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36.
14. Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions [Internet]. Version 5.1.0. London (England): The Cochrane Collaboration; 2011 Mar. Figure 15.5.a: Drummond checklist (Drummond 1996). Available from: http://handbook.cochrane.org/chapter_15/figure_15_5_a_drummond_checklist_drummond_1996.htm

15. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses [Internet]. Ottawa: The Ottawa Hospital Research Institute; 2014. [cited 2018 Apr 6]. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

16. Arbyn M, Verdoodt F, Snijders PJ, Verhoef VM, Suonio E, Dillner L, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. Lancet Oncol. 2014 Feb;15(2):172-83.

17. Gustavsson I, Aarnio R, Berggrund M, Hedlund-Lindberg J, Strand AS, Sanner K, et al. Randomised study shows that repeated self-sampling and HPV test has more than two-fold higher detection rate of women with CIN2+ histology than Pap smear cytology. Br J Cancer. 2018 Feb 13.

18. Williams DL, Hagensee .M., Gao R, Barnhill D, Fontham ETH. The accuracy and validity of HPV testing through self-collection with tampons for cervical cancer screening. Translational Cancer Research. 2016;5(Supplement5):S993-S999.

19. Lam JUH, Elfstrom KM, Ejegod DM, Pedersen H, Rygaard C, Rebolj M, et al. High-grade cervical intraepithelial neoplasia in human papillomavirus self-sampling of screening non-attenders. Br J Cancer [Internet]. 2018 Jan [cited 2018 Mar 22];118(1):138-44. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5765223/

20. Lam JUH, Rebolj M, Ejegod DM, Pedersen H, Rygaard C, Lynge E, et al. Prevalence of Human Papillomavirus in Self-Taken Samples from Screening Nonattenders. J Clin Microbiol [Internet]. 2017 Oct [cited 2018 Mar 22];55(10):2913-23. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5625377/

21. Ketelaars PJW, Bosgraaf RP, Siebers AG, Massuger LFAG, van der Linden JC, Wauters CAP, et al. High-risk human papillomavirus detection in self-sampling compared to physician-taken smear in a responder population of the Dutch cervical screening: Results of the VERA study. Prev Med. 2017 Aug [cited 2018 Mar 22];101:96-101.

22. Stanczuk G, Baxter G, Currie H, Lawrence J, Cuschieri K, Wilson A, et al. Clinical validation of hrHPV testing on vaginal and urine self-samples in primary cervical screening (cross-sectional results from the Papillomavirus Dumfries and Galloway-PaVDaG study). BMJ Open [Internet]. 2016 Apr 25 [cited 2018 Mar 22];6(4):e010660. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4854001/

23. Harvey LFB, Averbach SH, Hacker MR, Modest AM, Scott J, Ricciotti HA. Self-collection of vaginal swabs for human papillomavirus screening among women in temporary residential programs. Am J Obstet Gynecol [Internet]. 2016 Apr [cited 2018 Mar 22];214(4):546-7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4892360/

24. Haguenoer K, Giraudeau B, Gaudy-Graffin C, de P, I, Dubois F, Trignol-Viguier N, et al. Accuracy of dry vaginal self-sampling for detecting high-risk human papillomavirus infection in cervical cancer screening: a cross-sectional study. Gynecol Oncol. 2014 Aug;134(2):302-8.

25. Igidbashian S, Boveri S, Radice D, Casadio C, Spolti N, Sandri MT, et al. Performance of self-sampled HPV test in comparison with liquid based cytology. Eur J Obstet Gynecol Reprod Biol. 2014 Jun;177:72-6.

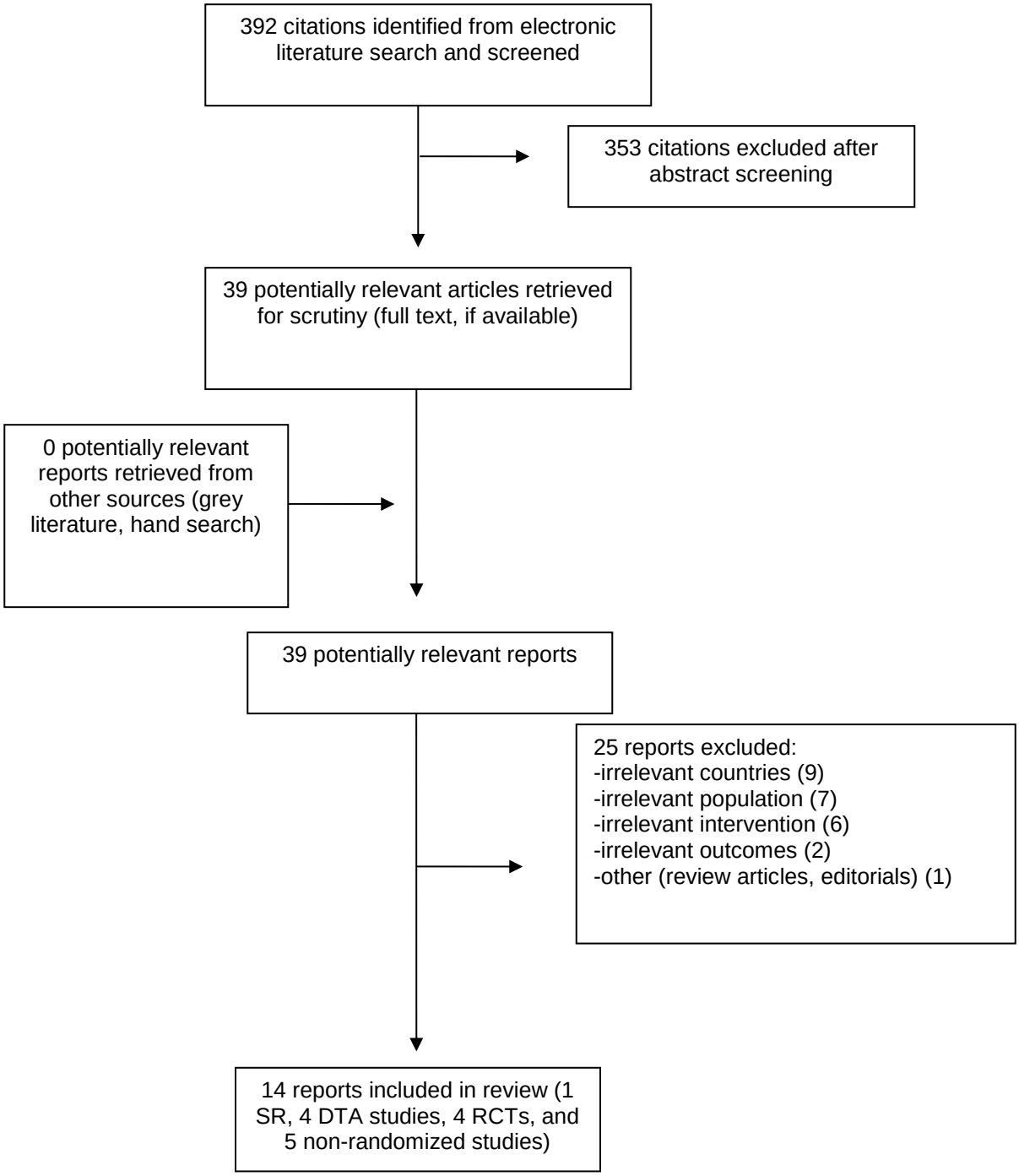
26. Jones HE, Mansukhani MM, Tong GX, Westhoff CL. Validity and reliability of using a self-lavaging device for cytology and HPV testing for cervical cancer screening: findings from a pilot study. PLoS ONE [Internet]. 2013 [cited 2018 Mar 22];8(12):82115. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3869665/

27. Darlin L, Borgfeldt C, Forslund O, Henic E, Hortlund M, Dillner J, et al. Comparison of use of vaginal HPV self-sampling and offering flexible appointments as strategies to reach long-term non-attending women in organized cervical screening. J Clin Virol. 2013 Sep;58(1):155-60.

28. Sancho-Garnier H, Tamalet C, Halfon P, Leandri FX, Le Retraite L, Djoufelkit K, et al. HPV self-sampling or the Pap-smear: a randomized study among cervical screening nonattenders from lower socioeconomic groups in France. Int J Cancer [Internet]. 2013 Dec 1 [cited 2018 Mar 22];133(11):2681-7. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.28283

29. Castle PE, Gage JC, Partridge EE, Rausa A, Gravitt PE, Scarinci IC. Human papillomavirus genotypes detected in clinician-collected and self-collected specimens from women living in the Mississippi Delta. BMC Infect Dis [Internet]. 2013 Jan 7 [cited 2018 Mar 22];13:5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3570306/

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 1: Characteristics of Included Systematic Reviews

First Author, Publication Year, Country	Types and Numbers of Primary Studies Included	Population Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes, Length of Follow-up
Arbyn et al. 2014,[16] Belgium	"Papers assessing the clinical accuracy of HPV DNA or RNA testing in self-samples" (34 studies included; 36 studies for data extraction and meta-analysis) 1 prospective cohort study, 1 retrospective cohort study, 2 RCTs, and 32 cross-sectional 16 screening studies, 17 follow-up studies, and 3 studies on high-risk populations. 1 follow-up article also using post-treatment samples.	154,556 women included. "16 studies (from 14 papers) were in primary screening of generally healthy women, whereas three studies screened high-risk populations. In 17 reports, women referred to colposcopy were enrolled. One study recruited women under follow-up and women who were treated for cervical precancer."	Self-sampled HPV testing: "a vaginal sample was self-taken by a woman (self-sample); a high-risk HPV DNA or RNA test was done or the clinician taken sample was examined microscopically for presence of cytological epithelial lesions" (p. 173)	Clinician-sampled HPV testing: "a sample taken by a clinician (clinician-taken sample) or self-samples were taken in one arm and clinician-taken samples in the other arm of randomised trials; a high-risk HPV DNA or RNA test was done or the clinician taken sample was examined microscopically for presence of cytological epithelial lesions" (p. 173) "In all but two selected studies, the comparator test was HPV testing on a clinician-taken sample, whereas in 20 reports, the clinician-taken samples were examined cytologically" (p. 176)	Sensitivities and specificities for the detection of CIN2+: "the presence or absence of CIN grade 2 (CIN2) or worse was verified by colposcopy and biopsy in all enrolled women or in women with at least one positive test."

CIN = cervical intraepithelial neoplasia; DTA = diagnostic test accuracy; HPV = human papillomavirus

Table 2: Characteristics of included primary studies

First Author, Publication Year, Country	Study design	Population Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes, Length of Follow-up
Lam et al. 2018,[19] Denmark	Non-randomized study on DTA, prospective cohort study	Non-attendees to the Copenhagen Self-sampling Initiative (CSi) pilot (n = 23,632) Mean age not reported	Self-sampled HPV tests using brushes; HPV tests including CLART, Onclarity, HC2 (	Cytology: population-based routine screening cohort from the Horizon study	DTA: positive predictive values for the detection of CIN2+; "Detection rate of ≥CIN2 among the screened women in a real-life self-sampling setting, the CSi,." (p. 139) 18-month follow-up since invitation
Stanczuk et al. 2016,[22] UK Papillomavirus Dumfries and Galloway (PaVDaG) study	Non-randomized study on DTA, prospective cohort study	"All women, other than those previously diagnosed with CIN2+, presenting for routine cervical screening (April 2013 to July 2014)" (p. 2) (n = 5318) Mean age = 41.3 years	Urine and self-sampled swab-based vaginal samples for HPV testing; Cobas HPV tests	Clinician-sampled HPV tests; LBC	DTA: "sensitivity and specificity of the assay in cervical and self-collected samples to detect CIN2+" (p. 3) A minimum of 8 months of follow-up
Igidbashian et al. 2014,[25] Italy	Non-randomized study on DTA, prospective cohort study	"All women scheduled for cervical cytology at our Institute" (p. 73) (n = 708) Mean age = 44.3 years	Self-sampled HPV tests using brushes; HC2 HPV tests	LBC	DTA: positive and negative predictive values for the detection of CIN2+ Follow-up length between 0.7 to 56.2 months
Jones et al. 2013,[26] USA	Non-randomized study on DTA, prospective cohort study	"Women who had attended one of three ambulatory clinics at the New York Presbyterian Hospital (NYPH) for cervical cancer screening." "Women needed to be at least 18 years old. Women were excluded if they reported current	Self-sampled HPV tests, lavage-based; HC2 HPV tests	LBC, clinician-sampled and lavage-based	DTAsensitivities and specificities for the detection of CIN2+; "Validity and reliability of using a self-lavaging device, the Delphi ScreenerTM (Delphi Bioscience, Scherpenzeel, Netherlands), for cervical cytology by comparing

First Author, Publication Year, Country	Study design	Population Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes, Length of Follow-up
		pregnancy, breastfeeding, hysterectomy, or discomfort reading on their own in Spanish or English." (p. e82116) (n = 198) Mean age of all participants not reported			paired self- and clinician-obtained specimens using LBC among 198 women, and high-risk HPV in a sub-sample" (p. e82115) 3 to 13 months of follow-up
Gustavsson et al. 2018,[17] Sweden	RCT	"The cervical cancer-screening programme in Uppsala County, Sweden, during that period invited women aged 23–49 years (3-year intervals) to Pap smear cytology testing, while women aged 50–60 years were invited to HPV testing (5-year intervals)." (p. 897) (n = 36,390) Mean age of all participants not reported Results presented in three age groups: 30 to 39, 40 to 49, and 30 to 49 years.	Self-sampled HPV tests using brushes: "a package including information on how to perform the sampling at home, a sampling brush, a FTA card and a preaddressed return envelope." (p. 897); RealTime PCR-based HPV tests	Cytology: "a midwife performed sampling on the cervix for Pap smear cytology" (p. 897)	Detection of CIN2+: "the number of women with CIN2+ based on histology diagnosed during the 18 months from date of invitation." (p. 897)
Williams et al. 2016,[18] USA	RCT	Participants living in underserved communities or those participating in a screening program for low-income, uninsured women aged 21 years of age and over had not had a	Self-sampled tampon-based HPV testing: "a kit which contained two tampons, a small pouch of personal lubricant, one plastic tube, one absorbent paper, one piece of bubble wrap, one biologics bag,	Clinician-sampled HPV testing: "appointment for a Pap test, HPV test, and pelvic exam in the gynecology clinic" (p. S995)	Agreement between self and clinician-sampled HPV testing : "Accuracy and validity of HPV testing with a sample self-collected with a tampon" (p. S994)

First Author, Publication Year, Country	Study design	Population Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes, Length of Follow-up
		hysterectomy, were English speaking, and were at least 1 year past their last Pap test (n = 120) Mean age of all participants not reported	an instruction sheet, a patient identification card, and a pre-addressed stamped postal service approved mailer" (p. S995); Cobas HPV tests		
Darlin et al. 2013,[27] Sweden	RCT	Women aged 32–65 years not taken a smear for nine years or more (n = 1,500) Mean age of all participants not reported	Self-sampled HPV testing using swabs; Luminex-based HPV	LBC	Acceptance and CIN2+ detection: "responses among long-term non-attending women to either (i) HPV-testing of a self-collected vaginal sample, or (ii) cytological screening with a flexible no-fee appointment for sampling at an outpatient clinic." (p. 156)
Sancho-Garnier et al. 2013,[28] France	RCT	"The women were living in the Bouches du Rhône area and had not had a Pap smear for 2 years. Within this group were all the women who had not responded to a first personal invitation to have a free of charge Pap-smear at a health center close to their home" (p. 2682) (n = 18,730) Mean age of all participants not reported	Self-sampled HPV testing using swabs: "a letter containing information about HPV infection and risk, the use of HPV detection for cervical cancer screening and they were notified that a self-sampling device for HPVHR detection would be sent to them soon" (p. 2682); RealTime HPV tests	Cytology: "invitations to have a Pap-smear" (p. 2682)	Detection of CIN2+: "rates of participation and rates of detection of CIN2" within 9 months (p. 2681)

First Author, Publication Year, Country	Study design	Population Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes, Length of Follow-up
		Participation also reported by five-year age groups			
Lam et al. 2017,[20] Denmark	Non-randomized study, prospective cohort study	Non-attendees to the Copenhagen Self-sampling Initiative (CSi) pilot, same population as Lam 2018[19] (n = 4801)	Self-sampled HPV tests using brushes; CLART, Onclarity, and HC2	Physician-sampled HPV testing: Population-based routine screening cohort from the Horizon study	Detection of HPV infections: "three assays (CLART, Onclarity, and HC2) and the subsequent cytology results for women with detected infections." (p. 2914)
Ketelaars et al. 2017,[21] the Netherlands	Non-randomized study, prospective cohort study	"Women aged 30–60 years and living in the regions of Nijmegen and 's-Hertogenbosch in the Netherlands, participated in the VERA study." (p. 97) (n = 2049)	Self-sampled HPV tests using brushes: "at the time of the appointment with their physician for their scheduled cervical smear, the participants also received a self-sampling kit including a self-sampling device" (p. 97); Cobas HPV tests	Cytology: "regular cervical smear taken by their physician as part of the nationwide program" (p. 97)	Detection of CIN2+: "concordance in high-risk HPV positivity between self-collected samples, using the Evalyn brush, and physician-taken samples" (p. 97) 9 to 16 months of follow-up
Harvey et al. 2016,[23] USA	Non-randomized study, cross-sectional study	"Women from 3 temporary residential programs. The programs include an emergency shelter for 26 families and 222-bed recovery programs at a Boston community health center" (p. 2) (n = 47)	Self-sampled HPV tests using swabs; HC2 HPV tests	Clinician-sampled HPV tests	Agreement between self and clinician-sampled HPV testing: "Accuracy and acceptability of self-collected vaginal swabs to detect high-risk HPV compared with physician-collected cervical swabs and cervical ThinPrep Papanicolaou tests" (p. 2)
Haguenoer et al. 2014,[24] France	Non-randomized study, cross-sectional study	"Women who were due for a routine screening Pap smear were	Self-sampled HPV tests using swabs; INNO-LiPA HPV tests	Clinician-sampled HPV tests with swabs	Agreement between self and clinician-sampled HPV testing:

First Author, Publication Year, Country	Study design	Population Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes, Length of Follow-up
		eligible if they were 20 to 65 years old, self-reported not a virgin, not pregnant, not vaccinated against HPV, not menstruating, had had no Pap smear for at least 2 years and had no prior hysterectomy" (p. 303) (n = 734)			"HR-HPV diagnostic accuracy of 2 vaginal self-collection methods and a clinician-collected sample which was also examined by cytology (i.e., Pap smear)." (p. 303)
Castle et al. 2013,[29] USA	Non-randomized study, cross-sectional	Women undergoing routine screening or had not been screened in the last three years. Women aged 26 to 65 years of age, non-pregnant, with a cervix, and willing to provide written, informed consent (n= 443)	Tampon-based self-sampled HPV testing; Linear Array HPV tests	Clinician-sampled HPV testing	Agreement between self and clinician-sampled HPV testing

CIN = cervical intraepithelial neoplasia; DTA = diagnostic test accuracy; HC2 = Hybrid Capture 2; HPV = human papillomavirus; LBC = liquid-based cytology; PCR = polymerase chain reaction; RCT = randomized controlled trial; VERA = Validation of the Evalyn brush with the Roche cobas 4800 hrHPV Test

Appendix 3: Critical Appraisal of Included Publications

Table 3: Strengths and Limitations of Systematic Reviews and Meta-Analyses using AMSTAR 2[12]

Strengths	Limitations
Arbyn et al. 2014,[16]	
• PICO components included in the research questions or study inclusion criteria • Study selection rationale described and explained • Reasons for study exclusion listed in the flowchart • Comprehensive search with PubMed, Embase and CENTRAL • Potential studies in all languages screened • Critical appraisal with QUADAS-2 checklist • Study selection in duplicate • Included studies described • A bivariate model used for meta-analysis of DTA • Heterogeneity due to country status, HPV sampling devices, and HPV tests discussed • Potential studies in all languages screened • Conflict of interest declared	• Protocol not established *a priori* • Data extraction not in duplicate • A list of excluded studies not provided • Funding sources of the included studies not mentioned • Risk of bias not explicitly tested in the meta-analysis • Risk of bias in individual studies not discussed regarding the results of meta-analysis

DTA = diagnostic test accuracy; QUADAS = quality assessment of diagnostic accuracy studies; PICO = population, intervention, comparator, and outcome

Table 4: Strengths and Limitations of DTA studies using QUADAS 2 checklist[13]

Strengths	Limitations
Lam et al. 2018[19]	
• Patient selection described • Case-control study avoided • Appropriate inclusion with routine screening populations • Index tests interpreted without the knowledge of reference standard • Diagnostic thresholds reported for HPV testing, except for HC2 • Comparator tests interpreted without the knowledge of reference standard • Diagnostic thresholds for the comparator tests, cytology, reported • Colposcopy used as reference standard • Appropriate intervals between initial screening tests and reference standard • Colposcopy as the only reference standard	• Colposcopy referral based on initial screening tests • Only test-positive patients after initial screening referred to reference standard • Only predictive values reported; not all patients analyzed for diagnostic test accuracy
Stanczuk et al. 2016[22]	
• Patient selection described and justified • Case-control study avoided • Appropriate inclusion with routine screening populations	• Colposcopy referral based on initial screening tests • Only test-positive patients after initial HPV and cytology tests referred to reference standard

Strengths	Limitations
• Index tests interpreted without the knowledge of reference standard • Diagnostic thresholds reported for HPV testing, Cobas • Comparator tests interpreted without the knowledge of reference standard • Diagnostic thresholds for the comparator tests, clinician-sampled HPV testing, reported • Colposcopy used as reference standard • Appropriate intervals between initial screening tests and reference standard • Colposcopy as the only reference standard • Verification bias considered and adjusted	
Igidbashian et al. 2014[25]	
• Patient selection described and justified • Case-control study avoided • Appropriate inclusion with routine screening populations • Index tests interpreted without the knowledge of reference standard • Diagnostic thresholds reported for HPV testing, HC2 • Comparator tests interpreted without the knowledge of reference standard • Diagnostic thresholds for the comparator tests, LBC, reported • Colposcopy used as reference standard • Appropriate intervals between initial screening tests and reference standard, within three years • Colposcopy as the only reference standard	• Colposcopy referral based on initial screening tests • Only test-positive patients after initial cytology tests referred to reference standard • Verification bias not adjusted
Jones et al. 2013[26]	
• Case-control study avoided • Appropriate inclusion with routine screening populations • Index tests interpreted without the knowledge of reference standard • Diagnostic thresholds reported for HPV testing, HC2 • Comparator tests interpreted without the knowledge of reference standard • Diagnostic thresholds for the comparator tests, LBC, reported • Colposcopy used as reference standard • Appropriate intervals between initial screening tests and reference standard, within 13 months • 10% test negative and test-positive patients after initial cytology tests referred to reference standard • Colposcopy as the only reference standard • Verification or ascertainment bias considered and adjusted	• A convenient sample recruited for study • Colposcopy referral based on initial screening tests

DTA = diagnostic test accuracy; HC2 = hybrid capture 2; HPV = human papillomavirus; LBC = liquid-based cytology

Table 5: Strengths and Limitations of RCTs using the Cochrane Risk of Bias checklist[14]

Strengths	Limitations
Gustavsson et al. 2018,[17]	
• Pathologists blinded • Attrition reported in Figure 2 • CIN2+ detection rates also resulting from attrition • Selective outcome reporting not likely	• Randomization method unclear • Allocation not concealed • Patients and physicians not blinded
Williams et al. 2016[18]	
• Attrition reported in Table 2 • CIN2+ detection rates also resulting from attrition • Selective outcome reporting not likely	• Randomization method unclear • Allocation not concealed • Pathologists not blinded • Patients and physicians not blinded • Clinician-sampled HPV tests as reference standard for self-sample tampon-based HPV tests
Darlin et al. 2013[27]	
• Attrition reported in Figure 2 • CIN2+ detection rates also resulting from attrition • Selective outcome reporting not likely	• Randomization method unclear • Allocation not concealed • Pathologists not blinded • Patients and physicians not blinded
Sancho-Garnier et al. 2013[28]	
• Computer-based randomization • Attrition reported in Figure 1 • CIN2+ detection rates also resulting from attrition • Selective outcome reporting not likely	• Allocation not concealed • Pathologists not blinded • Patients and physicians not blinded

CIN = cervical intraepithelial neoplasia; HPV = human papillomavirus; LBC = liquid-based cytology; RCT = randomized controlled trial

Table 6: Strengths and Limitations of non-randomized studies using Newcastle-Ottawa scale[15]

Strengths	Limitations
Lam et al. 2017[20]	
• Representing non-attendees in the Copenhagen Self-Sampling Initiative" (CSi) • Historical comparison with the Horizon study • Exposure documented in medical records • Outcome of interest, detection of HPV infection, not presenting at the start of the study • Control cohort sampled in the same region, Copenhagen • Outcome assessed with HPV assays • Cross-sectional study design appropriate for the outcome • Follow-up adequate for the assessment of test acceptance	• Differential response rates to self-sampled and physician-sampled HPV tests for the assessment of HPV infection
Ketelaars et al. 2017[21]	
• The same population receiving both the intervention and control in this co-testing study, self- versus clinician-sampled HPV testing • Exposure of intervention well documented • Intervention and comparator tested on the same population, although clinician-sampled HPV testing was required to be implemented first • Exposure to the intervention documented • Cross-sectional design sufficient for outcome assessment • Attrition reported	• Unclear representativeness of the study samples • Population inclusion and exclusion criteria unclear • Unclear about whether the outcome of interest, CIN2+, was part of the inclusion or exclusion criteria
Harvey et al. 2016[23]	
• The same population receiving both the intervention and control in this co-testing study, self- versus clinician-sampled HPV testing • Exposure of intervention well documented • Intervention and comparator tested on the same population • Exposure to the intervention documented • Cross-sectional design sufficient for outcome assessment • Attrition reported	• Unclear representativeness of the study samples • Population inclusion and exclusion criteria unclear • Unclear about the recruitment process or random sampling
Haguenoer et al. 2014,[24]	
• Representative of the women experiencing routine screening • Population inclusion and exclusion criteria clear • The same population receiving both the intervention and control in this co-testing study, self- versus clinician-sampled HPV testing • Exposure of intervention well documented • Intervention and comparator tested on the same population • Exposure to the intervention documented • Cross-sectional design sufficient for outcome assessment • Attrition reported	• Clinician-sampled HPV testing as reference standard

Strengths	Limitations
Castle et al. 2013,[29]	
• Somewhat representative of the women experiencing routine screening • The same population receiving both the intervention and control in this co-testing study, self- versus clinician-sampled HPV testing • Outcome of interest not present at the start of the study • Exposure to the intervention documented • Cross-sectional design sufficient for outcome assessment • Attrition reported	• The ratio of screened and under-screened women might not be applicable in other settings (252 versus 191)

CIN = cervical intraepithelial neoplasia; HPV = human papillomavirus

Appendix 4: Main Study Findings and Author's Conclusions

Table 7: Summary of Findings of Systematic Reviews

Main Study Findings	Author's Conclusions
Arbyn et al. 2014[16]	
• The studies on routine screening populations, high-risk groups, and follow-up were merged for the comparison of diagnostic test accuracy between self- and clinician-sampled HPV tests. • The pooled sensitivity of HPV testing on self-samples was lower than HPV testing on a clinician-taken sample (ratio 0·88 [95% CI 0·85–0·91] for CIN2 or worse and 0·89 [0·83–0·96] for CIN3 or worse). Also specificity was lower in self-samples versus clinician-taken samples (ratio 0·96 [0·95–0·97] for CIN2 or worse and 0·96 [0·93–0·99] for CIN3 or worse). • HPV testing with signal-based assays on self-samples was less sensitive and specific than testing on clinician-based samples. [HC2 (n = 18) and Cervista (n = 1)] • Aptima testing was less sensitive but not less specific in self-samples versus clinician-taken samples. (n = 1) • By contrast, some PCR-based HPV tests generally showed similar sensitivity on both self-samples and clinician-based samples. [GP5+/6+ PCR (n = 5), SPF10 PCR (n = 2), Abbott Real Time hrHPV test (n = 1), DNA chip (n = 1), modified GP5+/6+ PCR with Luminex reading (n = 1), or MALDI-TOF (n = 1)] • When HC2 was used (n = 18), self-sampling with any device showed lower sensitivity with differences being statistically significant for brushes, swabs, and tampons (n = 8, 7, and 1 respectively), while there was no significant difference for lavage (n = 2). • With GP5+/6+ PCR, the sensitivity of HPV testing on self samples collected with a brush or lavage device was similar to that seen with a clinician-collected brush sample • No differences by country status [high-income (n = 14) vs low-income or middle-income (n = 19)] could be discerned in the relative sensitivity. • HPV testing on self-samples was less sensitive and less specific than cytology with ASC-US or worse as a cut-off on clinician-taken samples with respect to detection of CIN2 or worse. However, for the detection of CIN3 or worse, HPV testing on self-samples was as sensitive as ASC-US or worse cytology on clinician specimen (table 3, appendix).	• In screening programmes using signal-based assays, sampling by a clinician should be recommended. • However, HPV testing on a self-sample can be suggested as an additional strategy to reach women not participating in the regular screening programme. • Some PCR-based HPV tests could be considered for routine screening after careful piloting assessing feasibility, logistics, population compliance, and costs.

ASC-US = atypical squamous cells of undetermined significance; CIN = cervical intraepithelial neoplasia; HPV = human papillomavirus; hr = high-risk; MALDI-TOF = matrix assisted laser desorption ionization-time of flight mass spectrometry; PCR = polymerase chain reaction

Table 8: Summary of Findings of DTA Studies

Main Study Findings	Author's Conclusions
Lam et al. 2018[19]	
• "Women participating in self-sampling had a higher CIN2+ detection than women undergoing routine cytology-based screening (OR = 1.83, 95% CI: 1.21–2.77) and a similar detection as routinely screened women tested with cytology and HPV testing (OR = 1.03, 95% CI: 0.75–1.40)." • "The positive predictive value for CIN2+ was higher in screening non-attenders than in routinely HPV- and cytology-screened screened women (36.5% vs 25.6%, respectively)." (p. 138)	• "Self-sampling offered to non-attenders showed higher detection rates for CIN2+ than routine cytology-based screening, and similar detection rates as HPV and cytology co-testing." (p. 138)
Stanczuk et al. 2016[22]	
• "Sensitivity for detecting CIN2+ was 97.7% (95% to 100%), 94.6% (90.7% to 98.5%) and 63.1% (54.6% to 71.7%) for cervical (clinician-sampled), vaginal (self-sampled) and urine hrHPV detection, respectively." • "The corresponding specificities were 87.3% (86.4% to 88.2%), 85.4% (84.4% to 86.3%) and 89.8% (89.0% to 90.7%)." • "There was a 38% (24% to 57%) higher HPV detection rate in vaginal self-samples from women over 50 years compared with those ≤29 years." (p. 1)	• "The sensitivity of self-collected vaginal (self-sampled) samples for the detection of CIN2+ was similar to that of cervical (clinician-sampled) samples" (p. 1) • LBC was less sensitive and more specific than clinical-sampled HPV tests.
Igidbashian et al. 2014[25]	
• "The sensitivity in detecting a CIN2 or worse resulted 0.4 for LBC and 0.8 for self-sampled HPV tests (with a positivity RLU cut-off ≥ 1). HPV test sensitivity was higher but not statistically different from the cytology sensitivity (*P* = 0.068)." • "The specificity in detecting a CIN2 or worse resulted 0.93 for the Pap smear and 0.81 for the self-sampled HPV tests." • "Cytology specificity was significantly higher than the specificity of the self-sampled HPV tests (*P* < 0.0001)." (p. 74)	• "Self-collected HPV testing identifies a group of women at high risk of positive LBC and high grade SIL." (p. 72)
Jones et al. 2013[26]	
• "Seven of 167 (4%) women with definitive results had CIN2+; one had normal and six abnormal cytology results with the self-lavage (sensitivity = 86%, 95% Confidence Interval, CI: 42, 100)." • "The kappa for paired cytology was low (0.36; 95% CI: 0.25, 0.47)" • "Seventy-three women had paired high-risk HPV tests with a kappa of 0.66 (95% CI: 0.49, 0.84)" (p. e82115)	• "A larger study to estimate the performance of the Screener for co-testing cytology and HPV or for HPV testing with cytology triage is warranted" (p. e82115)

ASC-US = atypical squamous cells of undetermined significance; CI = confidence interval; CIN = cervical intraepithelial neoplasia; DTA = diagnostic test accuracy; HPV = human papillomavirus; hr = high-risk; OR = odds ratio; PCR = polymerase chain reaction; RLU = relative light unit; SIL = squamous intraepithelial lesion

Table 9: Summary of Findings of RCTs

Main Study Findings	Author's Conclusions
Gustavsson et al. 2018,[17]	
• "Participation rate was 47% in the HPV arm and 39% in the control arm" • "For the per-protocol approach, cumulative prevalence of histological CIN2+ in the HPV arm was 20.2 per 1000 women screened as compared to 10.8 in the control arm." (p. 896)	• "Repeated self-sampling of VF and HPV test had more than a two-fold higher discovery rate of CIN2+ per 1000 women screened as compared with PAP smear cytology" (p. 896)
Williams et al. 2016[18]	
• **"**The percentage of tampon samples returned was 80.0% (48/60), which was significantly higher than attending of clinic visits (56.7%, 34/60). A valid HPV test was seen more often in clinician-obtained samples (34/34, 100%) as compared to self-collected samples (35/48, 72.9%).**"** • **"**Only 23 subjects in the Self-collection Arm had both the tampon sample and clinic sample collected, and HPV testing from the self-collected samples showed high sensitivity (100%) and specificity (94.1%) as compared to any HPV positive results with clinic sample as the reference test." (p. S993)	• "When collected and transported properly, self-collection with a tampon compared favorably to a clinical-obtained cervical swab for HPV testing." (p. S993)
Darlin et al. 2013[27]	
• "The response rate to HPV self-sampling was three times higher than the flexible outpatient clinic invitations (147/1000 women (14.7%) compared to 21/500 (4.2%) $p < 0.0001$)." (p. 155)	• "Although the response rate was low for both interventions, the invitation to vaginal HPV self-sampling was more effective for increasing the coverage of the screening programme." (p. 155)
Sancho-Garnier et al. 2013[28]	
• "Participation rates were significantly different between the two groups with only 2.0% of women attending for a Pap-smear while 18.3% of women returned a self-sample for HPV testing ($p \leq 0.001$). The detection rate of high-grade lesions (+CIN2) was 0.2& in the Pap-smear group and 1.25& in the self-sampling group ($p < 0.01$)." (p. 2681)	• "Offering self-sampling increased participation rates while the use of HPV testing increased the detection of cervical lesions (CIN2+) in comparison to the group of women receiving a second invitation for a Pap-smear. However, low compliance to follow-up in the self-sampling group reduces the effectiveness of this screening approach in non-attenders women and must be carefully managed." (p. 2681)

CIN = cervical intraepithelial neoplasia; DTA = diagnostic test accuracy; HPV = human papillomavirus; hr = high-risk; LBC = liquid-based cytology; OR = odds ratio; PCR = polymerase chain reaction; RLU = relative light unit; SIL = squamous intraepithelial lesion

Table 10: Summary of Findings of Non-Randomized Studies

Main Study Findings	Author's Conclusions
Lam et al. 2017[20]	
• "Non-attenders had an HPV prevalence of 11.3% as determined by the CLART assay, which was lower than that for women from the Horizon study (18.5%)." • Agreement between self- and clinician-sampled HPV testing: "One-third of the women who tested HPV positive by self-sampling tested HPV negative on the physician-taken follow-up sample." • "The CLART and Onclarity assays agreed on 64% (95% confidence interval [CI], 60 to 68%) of the HPV-positive self-taken samples. When the HC2 assay results were added into a three-way comparison, the level of agreement decreased to 27% (95% CI, 24 to 29%)." (p. 2913)	• "Further validation of HPV assays on self-taken samples is needed" (p. 2913)
Ketelaars et al. 2017[21]	
• "The hrHPV prevalence was 8.0% (95% CI 6.9–9.2) among the physician-taken samples, and 10.0% (95% CI 8.7–11.3) among the self-samples." • "There was 96.8% (95% CI 96.0–97.5) concordance of hrHPV prevalence between self-samples and physician-taken samples." • "Women in our study evaluated self-sampling as convenient (97.1%), user-friendly (98.5%), and 62.8% preferred self-sampling over a physician-taken sampling for the next screening round." (p. 96)	• "Self-sampling showed high concordance with physician-taken sampling for hrHPV detection in a responder screening population and highly acceptable to women." (p. 96)
Harvey et al. 2016[23]	
• "Using the physician-collected swab as the gold standard, the self-collected swab had a sensitivity of 84.6% (95% confidence interval [CI], 54.6–98.1%), specificity of 88.2% (95% CI, 72.6–96.7%), positive predictive value of 77.3% (95% CI, 44.9–92.2%), and negative predictive value of 93.8% (95% CI, 79.2–99.2%)." (p. 2)	• "Vaginal self-swab for HPV detection was a well-accepted and accurate method for cervical cancer screening." • "Although most participants found self-collection more private and easier, fewer women preferred self-collection. This apparent discrepancy may be due to concerns of incorrect self-collection, as informally reported by some participants" (p. 2)
Haguenoer et al. 2014,[24]	
• "Estimated sensitivity and specificity to detect HR-HPV in vsc-DRY samples were 88.7% and 92.5%, respectively, and in vsc-LIQ samples, 87.4% and 90.9%." (p. 302)	• "Vaginal self-sampling with a dry swab is accurate to detect HR-HPV infection as compared with cervical clinician-collection and accurate as compared with cytology results." (p. 302)
Castle et al. 2013,[29]	
• "The prevalence of carcinogenic HPV was 18.0% (95% CI: 14.4%-22.1%) for clinician-collected specimens and 26.8% (95% CI: 22.6%-31.4%) for self-collected specimens." • "The concordance for the detection of carcinogenic HPV between clinician-collected and self-collected specimens was only fair (kappa = 0.54)." (p. 1)	• "The high carriage of HPV infection, along with lack of participation in cervical cancer screening by some women, may contribute to the high cervical cancer burden in the region." (p. 1)

DTA = diagnostic test accuracy; HPV = human papillomavirus; hr = high-risk; LBC = liquid-based cytology; OR = odds ratio; PCR = polymerase chain reaction; RCT = randomized controlled trial; RLU = relative light unit; SIL = squamous intraepithelial lesion

Appendix 5: Additional References of Potential Interest

Non-systematic review

Snijders PJ, Verhoef VM, Arbyn M, Ogilvie G, Minozzi S, Banzi R, et al. High-risk HPV testing on self-sampled versus clinician-collected specimens: a review on the clinical accuracy and impact on population attendance in cervical cancer screening. Int J Cancer [Internet]. 2013 May 15 [cited 2018 Mar 22];132(10):2223-36. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.27790