



Cervical Cancer Screening with Human Papillomavirus Testing

Scoping Report

Schmucker C,¹ Kaier K,² Rücker G,² Motschall E,² Röbl-Mathieu M,³ Meerpohl JJ¹

¹Institute for Evidence in Medicine, Medical Center—University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany.

²Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Centre, University of Freiburg, Germany.

³Private practice for gynecology/obstetrics, Connollystr. 4, 80809 Munich, Germany.

Contact

Institute for Evidence in Medicine

Medical Center - University of Freiburg,
Faculty of Medicine, University of Freiburg,
Breisacher Str. 153, D -79110 Freiburg



Director of the Institute

Prof. Dr. med. Jörg Meerpohl, Mail: meerpohl@ifem.uni-freiburg.de

Senior Scientist (Project lead)

Dr. Christine Schmucker, Mail: schmucker@ifem.uni-freiburg.de

Content

Content	3
Abbreviations	6
1. Introduction	8
2. Objectives and research questions	9
3. Existing evidence syntheses	10
4. Methods	14
4.1. Search methods and study identification	14
4.1.1 Searches for published studies.....	15
4.1.2 Searches for unpublished and ongoing studies	15
4.1.3 Supplementary searches	15
4.1.4 Identification of relevant studies (screening process)	16
4.2 Clinical review	17
4.2.1 Participants/Population	17
4.2.2 Intervention (index test [screening test/strategy]).....	18
4.2.3 Comparator (comparator screening test/strategy)	18
4.2.4 Reference standard	19
4.2.5 Outcomes	19
4.2.6 Setting	21
4.2.7 Study types	21
4.2.8 Study duration.....	21
4.2.9 Search methods.....	22
4.2.10 Data extraction.....	22
4.2.11 Bias and GRADE assessment	23
Risk of bias	23
Dissemination bias (publication bias)	23
GRADE assessment	24

4.2.12	Data analyses.....	25
	Data from different study types	25
	Effect measures	25
	Assessment of heterogeneity	26
	Subgroup analyses	26
	Sensitivity analyses	26
	Unit of analysis.....	26
	Dealing with missing data	26
4.3	Review on health-related preferences	28
4.3.1	Participants/Population	28
4.3.2	Screening tests/strategies.....	28
4.3.3	Phenomena of interest (outcomes)	28
4.3.4	Setting	28
4.3.5	Study types	28
4.3.6	Study duration.....	29
4.3.7	Search methods.....	29
4.3.8	Data extraction.....	29
4.3.9	Quality assessment.....	30
4.3.10	Synthesis of findings	30
4.4	Economic review.....	32
4.4.1	Systematic review	32
	Participants/Population	32
	Screening tests/strategies and reference standard	32
	Outcomes.....	32
	Setting	33
	Study types	33
	Time horizon	33

Search methods	33
Data extraction	33
Quality assessment	34
Synthesis of findings	34
Assessment of the external validity/applicability to the Swiss setting.....	35
4.4.2 Cost-effectiveness analysis.....	35
Outcomes.....	36
Perspective	36
4.4.3 Budget impact analysis.....	37
4.4.4 Sensitivity analysis.....	37
5. References.....	39
Appendix A. Pilot search strategy.....	43

Abbreviations

ARTISTIC – Randomized Trial In Screening To Improve Cytology
ASCUS – Atypical Squamous Cells of Undetermined Significance
BIA – Budget Impact Analysis
CADTH – Canadian Agency for Drugs and Technologies in Health
CHEERS – Consolidated Health Economic Evaluation Reporting Standards
CI – Confidence Interval
CIN – Cervical Intraepithelial Neoplasia
GRADE – Grading of Recommendations, Assessment, Development and Evaluation
HIV – Human Immunodeficiency Virus
HPV – Human Papillomavirus
HR – Hazard Ratios
HTA – Health Technology Assessment
ICER – Incremental Cost-Effectiveness Ratio
ICTRP – International Clinical Trials Registry Platform
IHBR – Incremental Harm-Benefit Ratios
IQWiG – Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRR – Incidence Rate Ratios
ITT – Intention-to-Treat
LBC – Liquid Based Cytology
LSIL – Low Grade Squamous Intraepithelial Lesion
LY – Life-Year
MMAT – Mixed Methods Appraisal Tool
NHS EED – National Health Service Economic Evaluation Database
NRSs – Non-Randomized Studies
PICO – Patients, Intervention, Control, Outcomes
PRESS – Peer Review of Electronic Search Strategies
PRISMA-P – Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
QALY – Quality-Adjusted Life-Years

RCTs – Randomized Controlled Trials
RevMan – Review Manager
ROBINS-I: Risk of Bias in Non-Randomized Studies of Interventions
RoM – Ratio of Means
SMD – Standardised Mean Difference
SoF – Summary of Findings
SR – Systematic Review
UK – United Kingdom
US – United States
WHO – World Health Organization

1. Introduction

Cervical cancer remains a major public health problem worldwide, being the fourth most common cancer affecting women, with 570,000 new cases and 311,000 deaths in 2018.¹ Even though cervical cancer is one of the most preventable types of cancer known, and despite extensive vaccination and screening programs in most of the developed countries, women continue to suffer and die from it. In the period from 2006 to 2016, 878 women died from cervical cancer in Switzerland alone.²

Invasive cervical cancer generally develops over years, preceded by progressive changes of the cervical tissue, defined as cervical intraepithelial neoplasia (CIN) categorised as CIN 1, CIN 2, and CIN 3.³ Human papillomavirus (HPV) is the major risk factor for the development of cervical cancer, with approximately 99% of cervical cancers being associated with the virus.^{4 5} Furthermore, HPV can be detected with diagnostic tests that identify the presence of the virus. In view of the fact that HPV is the cause for the majority of cervical cancer cases, it is reasonable to explore further the idea of basing the screening programs on a HPV screening strategy rather than on conventional cytology—the long-standing screening recommendation with well-established benefits and harms regarding the procedure.

The Swiss Cancer Screening Committee commissioned a scoping report including description of planned methods for a full health technology assessment (HTA) on HPV screening approaches. Based on this report a decision will be taken whether to proceed with a full HTA to inform the decision whether HPV screening should replace cytology-based testing in Switzerland as the primary screening approach for cervical cancer, and what criteria, including appropriate screening intervals and ages to start and stop screening, should guide HPV-based cervical screening programs in Switzerland.

2. Objectives and research questions

The objective of this research work (HTA) is to evaluate evidence from studies on the comparative effectiveness and safety of screening approaches that use primary HPV screening in comparison to cytology-based screening in asymptomatic women.

The following four key questions will be addressed:

- What are the **benefits** of primary HPV screening, with or without cytology, compared with cytology-based testing in cervical cancer screening of asymptomatic women (addressed in section 4.2)?
- What are the **potential adverse effects (harms)** of primary HPV screening, with or without cytology, compared with cytology-based testing in cervical cancer screening of asymptomatic women (addressed in section 4.2)?
- What are the **health-related preferences** of asymptomatic women undergoing cervical cancer screening (primary HPV screening with or without cytology, and/or cytology-based testing (addressed in section 4.3)?
- What is the **cost-effectiveness** of primary HPV screening, with or without cytology, compared with cytology-based testing in cervical cancer screening of asymptomatic women (addressed in section 4.4)?

3. Existing evidence syntheses

A rapid scoping of existing evidence syntheses (i.e. HTAs, systematic reviews (SRs) and evidence-based guidelines) was conducted to inform the preparation of this scoping report. In total, eight evidence syntheses (including five HTAs, two systematic reviews, one IQWiG [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen] report)⁶⁻¹³ were identified and examined in more detail to determine their relevance to the question of this review. Key characteristics including the year when the report was published, country of publication, date of the last search for studies, number of included studies, age range of women, details regarding the screening tests and reference standard, collected outcomes and information whether a quality assessment of the studies was performed or not are summarized in Table 1.

We decided to capture data from international evidence syntheses published in the last five years (i.e., most up-to date reviews). More HTAs and/or reviews would be available, however, owing to a lacking currentness of data, we did not display their study characteristics. Three reports have been published in 2019.^{7 8 13} The latest one was published on the 10th of July 2019 from the Haute Autorité de Santé (France) and the last search for studies was conducted in January 2019. This report is based on 10 randomized controlled trials (RCTs) (addressing the efficacy of screening) and one meta-analysis (addressing diagnostic test accuracy). However, both harms of screening and women's preferences regarding different screening strategies were not (sufficiently) considered in this HTA. The recommendations of the United Kingdom (UK) National Screening Committee⁸ are solely based on one study from Greater Manchester providing a 10 year follow-up: ARTISTIC—Randomized Trial In Screening To Improve Cytology (<https://legacyscreening.phe.org.uk/cervicalcancer>). Cost-effectiveness was not considered in this research. The HTA from the Canadian Agency for Drugs and Technologies in Health (CADTH) is comprehensive including more than 20 studies and reviews.¹³ It is the only HTA addressing both comparative clinical and cost-effectiveness, and women's preferences on cervical cancer screening. It is also the only HTA including cytology-based testing with HPV triage as comparator screening test. This review includes studies published up to March 2017. The research question of the recommendations provided by the United States (US) Preventive Services Task Force is similar to ours.⁶ Cost-effectiveness and preferences were, however, not considered in this work. The last search for studies was performed early 2017 (similar to the Canadian HTA). The other HTAs or reviews include studies published before January 2016. In view of the rapidly expanding primary research (in terms of clinical studies) on HPV testing/screening, it is very likely that additional primary studies have been published since the end of the literature search of these

existing reports. Therefore, conducting a systematic review (HTA) based on the currently available primary research would be the most comprehensive approach to fully address the benefits, harms, women's preferences regarding screening and cost-effectiveness of HPV testing for cervical cancer screening.

Table 1. Other HTAs and systematic reviews evaluating HPV screening.

Year published	Country	Last search	Studies (N)	Asymptomatic women (age in years)	Screening test		Reference standard	Outcomes					Quality assessment
					Indextest	Comparator test		DTA (SE/SP)	Benefits	Harms	Women's preferences	Cost-effectiveness	
2019	France (HTA) ⁷	01/2019	RCTs (10) MA (1)	25-65	HPV testing with/without cytology	Cytology-based testing (conventional or liquid-based)	Colposcopy with biopsy when indicated	+	-Cervical cancer incidence and survival -Detection rate: CIN2+	-	-	+	+ Cochrane RoB ¹⁴ ; QUADAS-2 ¹⁵
2019	UK (HTA) ⁸	-	RCTs (1)	20-64	HPV testing with cytology	Cytology-based testing (liquid-based with HPV triage)	Colposcopy with biopsy when indicated	-	-Detection rate: CIN3+ -Screening intervals -Role of HPV genotyping	-	-	-	-
2019	Canada (HTA) ¹³	03/2017	SRs (4) RCTs (9) NRSs (11)	≥21	HPV testing with/without cytology	Cytology-based testing (with/without HPV triage)	Colposcopy with biopsy when indicated	+	-Cervical cancer incidence, morbidity and mortality	-FP and FN rates -Psychological effects -Adverse pregnancy outcomes -Any other harms	+\$	+	+ Cochrane RoB ¹⁴ ; AMSTAR-2 ¹⁶ ; QUADAS-2 ¹⁵ ; NOS ¹⁷
2018	USA (SR) ⁶	02/2017	RCTs (8) NRSs (5) IPD MA (1)	≥21	HPV testing with/without cytology	Cytology-based testing (conventional or liquid-based)	Colposcopy with biopsy when indicated	-	-Detection rate: CIN3+	-FP rates -Psychological effects	-	-	+ USPSTF ¹⁸ ; NOS ¹⁷
2017	Ireland (HTA) ⁹	01/2016	RCTs (7) NRSs (19) HTAs (3) CEAs (5)	18-70	HPV testing with/without cytology	Cytology-based testing (conventional or liquid-based)	Colposcopy with biopsy when indicated	+	-Detection rate: CIN2+/CIN3+	FP and FN rates	-	+	+ QUADAS-2 ¹⁵
2017	Cochrane (SR) ¹²	11/2015	RCTs (3) NRSs (37)	20-70	HPV testing with/without cytology	Cytology-based testing (conventional or liquid-based)	Colposcopy with biopsy when indicated	+	-	-	-	-	+ QUADAS-2 ¹⁵
2015	Belgium (HTA) ¹⁰	10/2013	RCTs (1) [§] NRSs (10)	18-70	HPV testing with/without cytology	Cytology-based testing (conventional or liquid-based)	Colposcopy with biopsy when indicated	+	-Detection rate: CIN2+/CIN3+	FP and FN rates	-	+	-
2014	Germany (IQWiG report) ¹¹	11/2013	RCTs (5)	20-60	HPV testing with/without cytology	Cytology-based testing (conventional or liquid-based)	Colposcopy with biopsy when indicated	+	-Detection rate: CIN2+/CIN3+	-	-	-	-

+ Aspect addressed in report; - Aspect not addressed in report.

AMSTAR-2: A Measurement Tool to Assess Systematic Reviews; CEA: Cost-Effectiveness-Analyses; CIN: Cervical Intraepithelial Neoplasia; DTA: Diagnostic Test Accuracy; FN: False Negatives; FP: False Positives; HTA: Health Technology Assessment; IPD-MA: Individual-Patient-Data Meta-Analyses; IQWiG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; MA: Meta-Analysis; N: Number; NOS: Newcastle-Ottawa Scale for non-randomized studies; NRS: Non Randomized Study; QUADAS-2: A revised tool for the Quality Assessment of Diagnostic Accuracy Studies; RCT: Randomized Controlled Trial; RoB: Risk of Bias; SE: Sensitivity; SP: Specificity; SR: Systematic Review; US: United States; USPSTF: US Preventive Services Task Force criteria.

* Sensitivity/specificity available only for subgroup analysis (HPV self-sampling vs professional sampling). The DTA results are based on a separate meta-analysis.¹⁹

§ Data on women's preferences regarding screening are based solely on qualitative studies.

§ This HTA describes the characteristics of 11 studies. However, for the DTA data, 60 studies were considered as relevant. The study flow is not clearly described in this HTA.

This IQWiG report is based on the update and the main report.

4. Methods

This scoping report has been conducted following the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P; <http://www.prisma-statement.org/Extensions/Protocols>) and will be registered in PROSPERO (International prospective register of systematic reviews).

The methods are subdivided into the following sections:

- Search methods and identification of studies (section 4.1);
- Methods for the effectiveness review addressing benefits and harms (section 4.2);
- Methods for the assessment of women's preferences (section 4.3);
- Methods for the cost-effectiveness assessment (section 4.4).

The methods described in this scoping report (protocol) will be followed throughout the full review process (HTA). Any deviations from the scoping report will be disclosed in the final report.

4.1. Search methods and study identification

Comprehensive systematic literature searches for relevant studies will be conducted by following the recommendation of PRESS (Peer Review of Electronic Search Strategies) and will be performed by an expert medical sciences librarian (EM).²⁰ The full electronic search strategies will be peer-reviewed by a second information specialist and validated by checking whether the strategy identified studies already known (e.g.,²¹⁻²⁶).

Search strategies for the databases mentioned below will be adapted from the (initial) Medline strategy (under construction, a first draft capturing the clinical review question is available in Appendix A). For the review addressing women's preferences the search strategy developed for the clinical review will be adapted and combined with additional search terms (also considering the operationalization of "preference" definitions). We will not apply study filters for different study designs because filters may not detect all studies that use the study designs (e.g., mixed-methods studies) considered for this research question (see section 4.3.5).²⁷

For the economic review the search strategy developed for the clinical review will be also used as "basis" and combined with economic search terms. Additionally, the NHS EED (National Health Service Economic Evaluation Database) filter will be applied when searching for economic studies in the database Medline and Embase.²⁸

We will not use any date or language restrictions in the electronic searches. For each database, the date of the search, the search strategy as well as the number of search results will be documented.

4.1.1 Searches for published studies

Searches for published studies will be conducted in the following electronic data sources:

- Medline, Medline Daily Update, Medline In Process & Other Non-Indexed Citations, Medline Epub Ahead of Print (via Ovid);
- Science Citation Index Expanded, Conference Proceedings Citation Index- Science Conference Proceedings Citation Index- Science, BIOSIS Citation Index (via Web of Science);
- Cochrane Library (via Wiley);
- Embase (via Embase.com/Elsevier).

To identify information related to women's preferences regarding HPV screening the following databases will be searched additionally to the ones stated above:

- PsycINFO (via Ovid);
- Cumulative Index to Nursing and Allied Health Literature (CINAHL; via EBSCO).

4.1.2 Searches for unpublished and ongoing studies

Searches for ongoing trials or unpublished completed studies will be performed in ClinicalTrials.gov (www.clinicaltrials.gov) and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (<http://www.who.int/ictrp/search/en>).

4.1.3 Supplementary searches

We will use relevant studies and/or systematic reviews to search for additional references via the Pubmed similar articles function (https://www.nlm.nih.gov/bsd/disted/pubmedtutorial/020_190.html) and forward citation tracking using the Web of Science Core Collection. Reference lists of relevant studies and systematic reviews will be reviewed and experts in the field (amongst others suggested by RM) will be contacted to enquire about any further relevant studies that may not have been retrieved by the electronic searches.

Furthermore, regular alerts will be established to update the searches until the final report is published and regular updates will be performed in databases that do not provide alert services.

Studies identified in the alerts that meet the selection criteria of the review will be incorporated into the analysis if they are identified prior to the completion of the stakeholder feedback period of the final report. Any studies that are identified after the stakeholder feedback period will be described in the discussion.

Clinical guidelines will be identified by searching different platforms such as <http://www.g-i-n.net> (Guidelines International Network), <https://www.tripdatabase.com> (Turning Research Into Practice), <http://www.sign.ac.uk> (Scottish *Intercollegiate Guidelines* Network), <https://www.nice.org.uk/guidance/published> (National Institute for Health and Care Excellence), <https://kce.fgov.be> (Belgian Health Care Knowledge Centre) or <http://www.oncoline.nl/index.php?language=en> (Cancer Clinical Practice Guidelines).

4.1.4 Identification of relevant studies (screening process)

Titles and abstracts of the citations identified by the searches will be independently screened by two reviewers (title and abstract screening [CS, NN]), and full texts of all potentially relevant articles will be obtained. Full texts will also be independently checked for eligibility by two reviewers (CS, NN), and reasons for exclusions will be documented (full text screening). Any disagreement will be resolved by consensus, moderated by a third reviewer. To standardize the screening process, two purpose-designed forms will be used: an 'abstract screening form' and a 'full text screening form'. Both will include the inclusion and exclusion criteria for the review. The 'abstract screening' form will be used as a guidance document to assist the title and abstract screening. The 'full text screening form' will be used to extract and document the key criteria that are relevant for the ultimate decision of whether a study will be included or excluded as well as the key justification for the decision by each reviewer. The title and abstract screening will be piloted on a random subset of 50 search results. The 'full text screening' will be piloted on three exemplary RCTs. The complete screening process will be conducted in Covidence (<https://www.covidence.org/home>).

4.2 Clinical review

4.2.1 Participants/Population

Inclusion criteria:

- Asymptomatic women;
- Women close to or within the age range suitable for cervical cancer screening according to the guidelines of the Swiss Society for Gynecology and Obstetrics (20 to 70 years).

Exclusion criteria:

- Women at high risk for cervical cancer (e.g., immunocompromised, HIV [human immunodeficiency virus]-positive);
- Women with cytological abnormalities or known cervical cancer;
- Women followed up for earlier cytological abnormalities;
- Women who received previous treatment for high-grade squamous intraepithelial lesions;
- Total or radical hysterectomy;
- Pregnant women.

Subgroups:

- Age (< 30 years vs ≥ 30 years);
- Vaccination status (i.e., HPV-vaccinated, stratified by vaccine type [i.e., bi-valent, quadri-valent, or nine-valent] vs not HPV-vaccinated);
- Screening history (e.g., none vs regular screening participation);
- Other relevant patient characteristics that stratify health-related outcomes (e.g., the socioeconomic status, number of pregnancies, smoker, number of sexual partners or ethnicity).

4.2.2 Intervention (index test [screening test/strategy])

We will consider the following screening strategies (see also Figure 1):

- HPV testing without cytology-based testing;
- HPV testing in combination with cytology (either conventional cytology [Pap smear] or liquid-based cytology [LBC]);
- HPV with cytology-based triage (either conventional cytology or LBC);
- Any screening interval;
- Any method of sample collection;
- HPV tests have to be currently approved by national and/or international health authorities and have to fulfil the Meijer criteria,²⁹ examples of eligible HPV tests include:
 - PCR using primers such as GP5+/GP6+, MY09/11, SPF10, CPI/II;
 - HC2 or newer improved signal amplification methods;
 - Aptima HPV assay, Aptima HPV 16 18/45 genotype assay;
 - Cervista HPV HR assay.

Subgroups:

- Screening interval (e.g., every year, every two years, every three years, every five years);
- Method of sample collection (self-collected vs clinician collected);
- Type of assay (e.g., generic, partial genotyping, or full genotyping);
- Type of HPV test (test category, e.g., PCR, HC2, Aptima);
- HPV test threshold for a positive result (e.g., for the HC2 method: 1 pg/mL vs 2 pg/mL).

4.2.3 Comparator (comparator screening test/strategy)

We will consider the following comparator screening strategies (see also Figure 1):

- Cytology-based testing (either conventional cytology or LBC) without HPV triage;
- Cytology-based testing (either conventional cytology or LBC) with HPV triage;
- Any screening interval will be considered.

Subgroups:

- Screening interval (e.g., every year, every two years, every three years, every five years);
- Cytology test threshold for a positive result (e.g., for LBC and conventional cytology we will consider two thresholds that define a positive result: ASCUS [atypical squamous cells of undetermined significance] or worse and LSIL [low grade squamous intraepithelial lesion] or worse, respectively).³⁰

4.2.4 Reference standard

The following reference standard will be considered:

- Colposcopy with histological examination of tissue specimens;
- Colposcopy without histological examination of tissue specimens;
- Reference standard needs to be applied to:
 - All women; or
 - All women with a positive screening test and a subset of screening test-negative women; or
 - All women with a positive screening test.

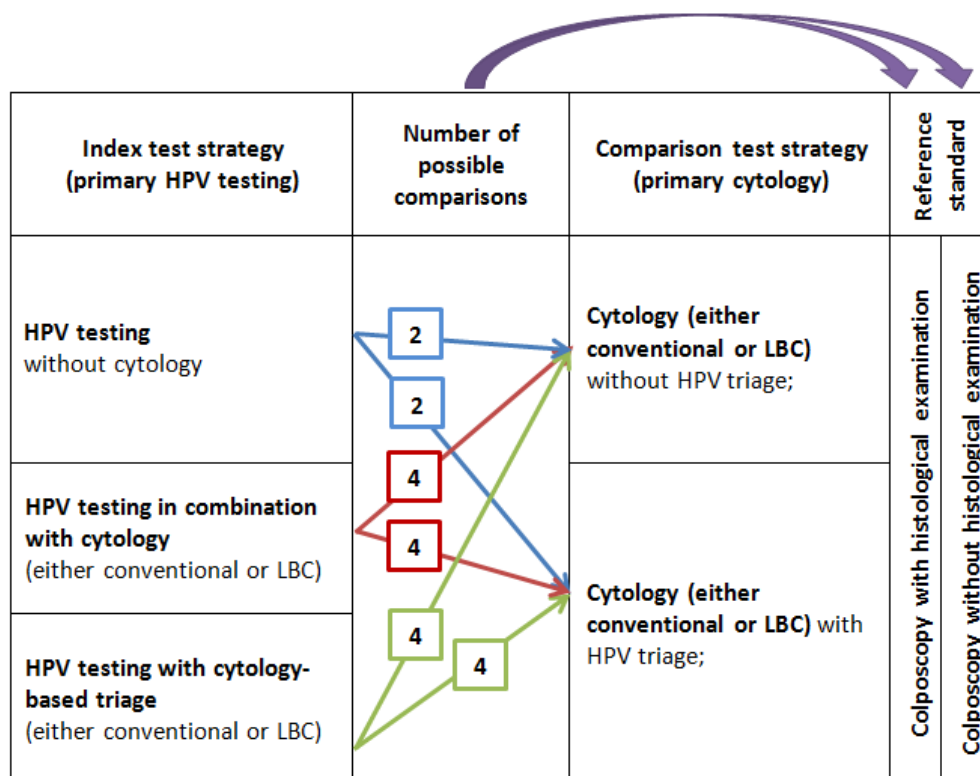


Figure 1. Displays the different screening test strategies considering different combinations of the index test (primary HPV testing), different combinations of the comparison test (primary cytology-based testing) and two options of the reference standard (colposcopy with and without a histological examination). Possible subgroup analyses such as different screening intervals, different sampling methods, different age groups or vaccination status are not considered in Figure 1. If we identify studies not differing between the type of cytology used (conventional or LBC), we will categorize them as “cytology (not further defined)”. Of note, this possible stratification is not displayed in Figure 1.

4.2.5 Outcomes

Patient-relevant outcomes in terms of benefits and harms:

Benefits

- Mortality:
 - Overall survival;
 - Cervical cancer-related mortality.
- Morbidity (in terms of cumulative incidence and incidence density [at the study level]):
 - Incidence of precursor lesions of cervical cancer (CIN 2/CIN 2+, CIN 3/CIN 3+,);
 - Incidence of cervical cancer overall (invasive cancer; categorized by different stages [e.g., stage IIA1 and IIA2]) and categorized by histological type (adenocarcinoma in situ, adenocarcinoma, squamous cell carcinoma)
 - Incidence of glandular lesions;
 - Incidence of endometrial cancer.
- Quality of life, as measured by standardized scales;
- Number of women referred to colposcopy (with and without histological examination);
- Number of women referred for treatment;
- Number of women accepting screening;
- Adherence.

Harms

- False-positive screening results: defined as the number of women with a positive screening test result without having precursor lesions or invasive cervical cancer; (direct harm of screening: overdiagnosis and/or unnecessary treatment and related impact on patients);
- False-negative screening results defined as the number of (precursor lesions or invasive cervical cancer) cases occurring among women with negative screening results; (direct harm of screening: undertreatment);
- Psychological harms (e.g., anxiety, labeling, stigma, distress);
- Adverse pregnancy outcomes;
- Adverse treatment effects;
- Any other reported harm.

4.2.6 Setting

Inclusion criteria:

- Outpatient screening within a screening program or opportunistic screening (including screening in family planning clinics) conducted in industrialized countries (European Economic Area countries, UK, New Zealand, Australia, US, Canada).

Exclusion criteria:

- Studies providing results from low-resource countries will be excluded due to lower applicability to the Swiss context.

4.2.7 Study types

Inclusion criteria:

- RCTs with randomization either at the individual or cluster level;
- Non-randomized studies (NRSs) including both (i) studies in which women (individuals or clusters of individuals) are allocated to different screening test strategies using methods that are not random; and (ii) observational studies, i.e., prospective and retrospective cohort studies using a longitudinal or cross-sectional design. In observational studies the allocation to the screening test/strategies is not determined by the study investigators, but by the nature of other factors outside the control of the investigator.

Exclusion criteria:

- Case-control studies;
- Case reports;
- Case series;
- Review articles;
- Work that has not been peer-reviewed (e.g., thesis, editorials, letters, comments);
- Results reported in an abstract form only (i.e., conference abstracts) will not be considered owing to limited information on study methods.

4.2.8 Study duration

We will not apply any exclusion criteria regarding study duration.

4.2.9 Search methods

The search methods are addressed in detail for each review (clinical review, review on health-related preferences, and economic review) in section 4.1.

4.2.10 Data extraction

Two review authors will extract the following study data and tabulate all relevant information:

- *Study characteristics*, i.e., author, year of publication, study type, start and end of study, sample size (total and for each study arm), follow-up time, inclusion/exclusion criteria, recruitment method, founding sources;
- *Setting*, i.e., geographical and organizational setting, e.g., screening within a screening program, opportunistic screening including health care provider (e.g., family medicine, obstetrics/gynecology, university-based health clinics, mobile clinics, family planning clinics, worksites);
- *Characteristics of the participants*, e.g., age range, vaccination status, screening history, socioeconomic status, ethnicity, education, number of pregnancies, smoking status, history of sexual activity, relationship status;
- *Characteristics of the intervention (index test)*, e.g., details on the screening strategy (e.g., HPV testing with/without cytology, number of rounds of screening), type of HPV test, definition for a positive screening result, method of sample collection (physician-collected, self-collected), screening interval;
- *Characteristics of the comparator*, e.g., details on the screening strategy (cytology-based testing with/without HPV triage, number of rounds of screening), type of cytology (conventional or LBC), definition for a positive screening result, screening interval;
- *Characteristics of the reference standard*, e.g., colposcopy (with and without histological examination), number of women to whom test was applied (e.g., all women, only screening test-positive women);
- *Time between screening test results and reference standard (if applicable)*;
- *Non-attendees*, number and reasons;
- *Outcome measures*, i.e., test description, data for continuous-, dichotomous-, or categorical-efficacy variables, unit of measurement, upper and lower scale limits, collected and reported time points of measurement. Where adjusted analyses are available in primary studies, these adjusted estimates will be used. Where adjusted analyses are not available, we will extract the unadjusted data as reported in the study.

Two reviewers (CS, NN) will pilot data extraction forms on three exemplary studies. Data from each included study will then be extracted by two reviewers (CS, NN) independently. Disagreements will be resolved through discussion until consensus is reached, involving a third reviewer if necessary. Moreover, authors of the studies included in the HTA will be contacted to provide any missing information or clarify any issues.

4.2.11 Bias and GRADE assessment

Risk of bias

Risk of bias of RCTs will be assessed according to the methodology described in the Cochrane Handbook for Systematic Reviews of Interventions.^{14 31} The following domains will be addressed: (i) bias arising from the randomization process; (ii) bias due to deviations from intended interventions; (iii) bias due to missing outcome data; (iv) bias in measurement of the outcome; and (v) bias in selection of the reported result. These domains will be judged with 'low risk of bias', 'some concerns' or 'high risk of bias'.

Bias in NRSs will be evaluated separately according to the 'Risk of Bias in Non-randomized Studies of Interventions' (ROBINS-I) tool, addressing the following domains: (i) bias due to confounding (e.g., age, screening history, vaccination status, socioeconomic differences); (ii) bias in selection of participants into the study (e.g. inception bias); (iii) bias in measurement of the intervention; (iv) bias due to departures from intended interventions; (v) bias due to missing data; (vi) bias in measurement of outcomes; (vii) bias in selection of the reported result; and (viii) overall bias.^{32 33} These domains will be judged as 'low', 'moderate', 'serious', 'critical' or 'unclear' risk of bias.

The ratings for the individual domains for each RCT and NRS will be presented in a 'risk of bias' table, separately.

Dissemination bias (publication bias)

We plan to minimize the impact of dissemination bias by ensuring a comprehensive search for eligible studies including searches of trial registries (see above 'search methods'). A funnel plot and appropriate statistical tests for small study effects will be performed if ≥ 10 studies are available addressing the same outcome.

GRADE assessment

The certainty of evidence of selected patient-relevant outcomes will be assessed using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach.³⁴

³⁵ The following prioritized outcomes (confirmed by the stakeholders) will be considered:

- Mortality:
 - Overall survival;
 - Cervical cancer-related mortality.
- Morbidity (in terms of cumulative incidence and incidence density [at the study level]):
 - Incidence of precursor lesions of cervical cancer (CIN 2/CIN 2+, CIN 3/CIN 3+,);
 - Incidence of cervical cancer overall (invasive cancer; categorized by different stages [e.g., stage IIA1 and IIA2]) and categorized by histological type (adenocarcinoma in situ, adenocarcinoma, squamous cell carcinoma)
 - Incidence of glandular lesions;
 - Incidence of endometrial cancer.
- Quality of life, as measured by standardized scales.
- Number of women referred to colposcopy (with and without histological examination);
- Number of women referred for treatment;
- False-positive screening results;
- False-negative screening results.

In brief, the GRADE assessment will address different aspects including:

- Study limitations (risk of bias);
- Imprecision (when 95% confidence intervals are wide and/or are close to null effect around the point estimate);
- Inconsistency (that is, differences in effect estimates across studies that assessed the same comparison);
- Indirectness (that is, differences in patient characteristics, differing [co-] intervention [testing], differing extent to which the intervention of interest is optimally conducted, differing comparator, and differences in measurement of outcome);
- Dissemination bias;
- Other potential criteria (e.g. large effect estimates) that can increase certainty.

Based on these criteria, the certainty of the evidence for each outcome can be categorized as either high, moderate, low, or very low. The results will be presented in a Summary of Findings (SoF) Table as suggested by the GRADE Working Group.^{36 37}

Risk of bias assessment and the GRADE assessment will be conducted independently by two reviewers (CS, NN). Any disagreements will be resolved by discussion and consensus involving, when needed a third person.

4.2.12 Data analyses

Data from different study types

RCT and NRS data will be analyzed—due to different mechanisms of bias—separately.^{9 38}

Effect measures

We intend to conduct meta-analyses for the comparisons defined in Figure 1 considering all predefined outcomes. The realization of the meta-analyses will depend on the availability of sufficient data from sufficiently homogenous studies in terms of clinical (e.g., age at screening, type of screening test, threshold for a positive finding) and methodological (considering the study design) homogeneity. Meta-analysis of data from NRSs will only be considered among studies with similar design (e.g. prospective cohort studies will only be combined with other prospective cohort studies). Meta-analyses will be conducted using Review Manager (RevMan) Version 5.3 and effect estimates will be calculated based on a random effects model.³⁹

Continuous outcomes: We will analyze outcomes measured with a scale as continuous outcomes (e.g. quality of life, anxiety). If more than one scale is used to measure an outcome in the very same study, only the measurement that has been obtained using a validated scale will be considered. In the case when more than one validated or more than one but only scales that are not validated are reported for one outcome, we will use the results provided by the scale that is most commonly used in the other included studies. The effect estimate for each continuous outcome will be expressed as the mean difference with its 95% confidence interval (CI). Where continuous outcomes are measured using different scales, the effect of the intervention will be expressed as the standardized mean difference (SMD) with its 95% CI. As recommended by Guyatt et al., and where possible, effect measures will additionally be expressed by the ratio of means (RoM) with their 95% CIs to facilitate interpretation.⁴⁰

Dichotomous outcomes: The effect estimate for dichotomous outcomes (e.g. mortality, morbidity) will be summarized using relative risks and 95% CI.

Time-to-event data: For outcomes reported as time-to-event, and given available individual patient data in the form of a survival curve or table of events per patients at risk, we will combine

time-to-event data using hazard ratios (HR) and incidence rate data using incidence rate ratios (IRR).

Assessment of heterogeneity

Different types of heterogeneity (owing to different clinical characteristics, different study designs or small study effects) will be evaluated and statistically quantified based on I^2 and the statistical test chi square. Thereby, an $I^2 \geq 75\%$ will be considered as considerable heterogeneity.⁴¹ Sensitivity analyses (considering the risk of bias) and predefined subgroup analyses (for different clinical characteristics including different populations, interventions and comparator) will be performed irrespective of the presence of "significant" heterogeneity (see below).

Subgroup analyses

Subgroup analyses will be performed to examine whether effect estimates will be affected by:

- Characteristics of the population (see suggested analyses, section 4.2.1);
- Characteristics of the screening test (indextest; see suggested analyses, section 4.2.2);
- Characteristics of the comparator (see suggested analyses, section 4.2.3).

Differences in effect estimates will be assessed by interaction tests available within RevMan Version 5.3. Furthermore, if a sufficient number of studies and data are available, meta-regression will be considered.

Sensitivity analyses

Sensitivity analyses will be conducted to determine the impact of bias through the exclusion of studies with high risk of bias. In case of any differences between these estimates, these will be considered in the results and discussion.

Unit of analysis

The unit of analysis will be the individual woman. For cluster-randomized studies we will adjust for clustering.⁴²

Dealing with missing data

Data will be analyzed—if possible—on intention-to-treat (ITT) basis or according to recently developed recommendations for systematic reviewers for addressing missing data in clinical

trials.^{43 44} We will also check trial register records or attempt to contact study authors to obtain information on missing data. If results are only reported in graphs, we will estimate the values based on these graphs.

In general, if pooling is not appropriate, a narrative synthesis will include the presentation of findings within summary tables alongside study and clinical characteristics believed to contribute to heterogeneity. A narrative description will aim to synthesize the direction and size of any observed effects across studies in the absence of a meta-analysis and will include an assessment of the likelihood of clinical benefit or harm.

4.3 Review on health-related preferences

The purpose of this systematic review will be to (i) explore current research addressing health-related preferences in asymptomatic women eligible for HPV screening and (ii) to synthesize these studies using quantitative and qualitative methods to support decision-making using an iterative process.

4.3.1 Participants/Population

The same inclusion and exclusion criteria as described in the methods of the clinical review will be applied (see section 4.2.1).

4.3.2 Screening test strategies

The same inclusion and exclusion criteria as described in the methods of the clinical review will be applied (see section 4.2.2 and 4.2.3).

4.3.3 Phenomena of interest (outcomes)

Our phenomena of interest will be:

- Preferences related to the organization of healthcare;
- Preferences related to desired, undesired and competing outcomes (in terms of safety and effectiveness of the screening test and screening strategy);
- Screening strategy preferences;
- Treatment preferences (after a positive result).

The classification and/or definition of the outcomes (type of preferences) will be discussed and consecutively adapted, depending on the literature findings.

4.3.4 Setting

The same inclusion and exclusion criteria as described in the methods of the clinical review will be applied (see section 4.2.6).

4.3.5 Study types

Inclusion criteria:

- We will include qualitative, quantitative (observational or interventional), or mixed methods studies (including both qualitative and quantitative data) that address health-

related preferences (including priorities, goal oriented, shared decision-making, patient centered, patient oriented, values, "satisfaction", "experiences", and "attitudes" toward HPV testing) from the women's perspective.

Exclusion criteria:

- Case reports;
- Review articles;
- Work that has not been peer-reviewed (e.g., thesis documents, editorials, letters, comments);
- Results reported in an abstract form only (i.e., conference abstracts) will also not be considered owing to limited information on study methods;
- Studies addressing the preferences of caregivers, family members and healthcare professionals will not be considered.

4.3.6 Study duration

We will not apply any exclusion criteria regarding study duration.

4.3.7 Search methods

The search methods are addressed in detail for each review (clinical review, review on health-related preferences, and economic review) in section 4.1.

4.3.8 Data extraction

Two review authors will extract the following study data and tabulate all relevant information:

- *Study characteristics*, i.e., author, year of publication, study design (qualitative, quantitative [observational or interventional], or mixed methods studies (including both qualitative and quantitative data), start and end of study, sample size, follow-up time, inclusion/exclusion criteria, recruitment method, founding sources;
- *Setting*, i.e., geographical and organizational setting, e.g., screening within a screening program, opportunistic screening including health care provider (e.g., family medicine, obstetrics/gynecology, university-based health clinics, mobile clinics, family planning clinics, worksites);
- *Aim of study*;

- *Characteristics of the screening population*, e.g., age range, vaccination status, screening history, socioeconomic status, ethnicity, education, number of pregnancies, smoking status, history of sexual activity, relationship status;
- *Characteristics of the screening tests*, i.e., characteristics of the index test (HPV screening) and, if applicable (depending on the design of the study), of the comparator test (cytology-based screening);
- *Characteristics of the reference standard* (if applicable, depending on the design of the study);
- *Non-attendees*, number and reasons;
- Preference (data) collection methods, i.e., methods used to elicit women's preferences (e.g., by, interviews, focus groups);
- *Phenomena of interest (outcome)*, i.e., type/definition of preferences (e.g., desired, undesired and competing outcome preferences, screening strategy preferences—as guided by the above description of the phenomena of interest) and study results.

Two reviewers (CS, NN) will pilot data extraction forms on three exemplary studies. Data from each included study will then be extracted by two reviewers (CS, NN) independently. Disagreements will be resolved through discussion until consensus is reached, involving a third reviewer if necessary. Moreover, authors of the studies included in the HTA will be contacted to provide any missing information or clarify any issues.

4.3.9 Quality assessment

We will appraise the risk of bias and applicability of the results of included studies using the Mixed Methods Appraisal Tool (MMAT) for Evaluating Primary Research Studies which allows the quality of qualitative, quantitative and mixed methods studies to be critically appraised.⁴⁵ Two reviewers (CS, NN) will assess study quality. Disagreement will be solved in discussion with a third reviewer.

4.3.10 Synthesis of findings

We will summarize the available data using tables and figures (i.e., bubble plots) to present the study landscape and to elicit preference clusters and/or gaps. We will describe the identified study pool considering characteristics such as study design, location, setting, population and subpopulations (e.g., considering age, vaccination status and others, if

possible), different HPV and/or cytology-based screening strategies, methods used to elicit women's preferences and the phenomena of interest (i.e., type of preference). The reporting of the results and data synthesis will follow an iterative process depending on the data availability. All members of the research team including external experts with focus on patient-centered care will be consulted during all steps of the review process.

When the included studies permit quantitative and qualitative information synthesis, we intend to conduct a mixed-methods systematic review using a convergent integrated approach that (i) synthesizes qualitative data, (ii) synthesizes quantitative data, and (iii) synthesizes both (i) and (ii).^{46 47}

4.4 Economic review

The health economic assessment (economic review) will include:

- (i) a systematic review of the currently published studies (section 4.4.1);
- (ii) a cost-effectiveness analysis (section 4.4.2);
- (iii) a budget impact analysis (BIA) (section 4.4.3).

4.4.1 Systematic review

Participants/Population

The same inclusion and exclusion criteria as described in the methods of the clinical review will be applied (see section 4.2.1).

Screening test strategies and reference standard

The same screening test strategies as described in the methods of the clinical review will be applied (see section 4.2.2, 4.2.3 and 4.2.4).

Outcomes

Mortality

- Overall survival;
- Cervical cancer-related mortality;
- LY gained.

Morbidity

- QALY gained;
- Prevented cancer cases (invasive cancer) and prevented cases of precursor lesions (e.g., CIN 2/CIN 2+, CIN 3/CIN 3+, glandular lesions).

Harms

- False-positive screening results (overdiagnosis and/or unnecessary treatment);
- False-negative screening results (undertreatment);
- Psychological harms (e.g., anxiety, labeling, stigma, distress);
- Adverse pregnancy outcomes;
- Adverse treatment effects (including toxicity);
- Any other reported harm.

Costs

- Costs per LY gained;
- Costs per QALY gained;
- Costs of different HPV screening strategies;
- Costs of different cytology-based screening strategies.

Outcomes will be stratified after different HPV screening strategies (e.g., considering different tests and test combinations, methods of sampling collection and screening intervals), the women's age and the vaccination status.

Setting

The same inclusion and exclusion criteria as described in the methods of the clinical review will be applied (see section 4.2.6).

Study types

We will include all types of economic evaluations studies including full economic evaluations (i.e., cost-benefit analysis, cost-effectiveness analysis and cost-utility analysis), partial economic evaluations studies and single effectiveness studies (e.g., RCTs, NRs, observational studies) reporting a full-scale incremental cost-effectiveness analysis (ideally, but not necessarily with an endpoint of costs per QALY gained or costs per LY gained).⁴⁸

Time horizon

The time horizon of the modelling approaches should provide us with both the short- and long term (ideally life-long) impact of different HPV screening strategies (including different screening test and screening intervals) on the expectation of life and quality of life.

Search methods

The search methods are addressed in detail for each review (clinical review, review on health-related preferences, and economic review) in section 4.1.

Data extraction

We will extract the following data from eligible economic evaluations studies:

- *Study characteristics*, e.g., author, year of publication, study design, start and end of study, sample size (total and for each study arm), follow-up time;

- *Setting*, e.g., geographical and organizational setting (e.g., screening within a screening program, opportunistic screening);
- *Characteristics of the participants*, e.g., age, screening history, socioeconomic status;
- *Characteristics of the intervention (index test)*, e.g., details on the screening strategy (e.g., HPV testing with/without cytology), type of HPV test, screening interval;
- *Characteristics of the comparator*, e.g., details on the screening strategy (e.g., cytology-based testing with/without HPV triage), type of cytology, screening interval;
- *Characteristics of the reference standard*, e.g., colposcopy (with and without histological examination);
- *Non-attendees*, number and reasons; and coverage rate;
- *Outcome measures*, e.g., different cost types/data, cost per year, QALY, LY, incremental cost-effectiveness ratio (ICER), effectiveness data (mortality, morbidity, harm);
- *Type and method of the cost-effectiveness model*;
- *Utility parameters/scores*;
- *Discount rate*;
- *Approaches to sensitivity analysis*;
- *Time horizon for both costs and effects*.

Quality assessment

Before conducting the quality assessment the included health economics studies will be classified by study design. Studies will then be assessed using the methods recommended by the Campbell and Cochrane Economic Methods Group⁴⁸; i.e., **critical appraisal** of the methodological quality of the economic modelling studies will be conducted using the “Phillips checklist”.^{49 50} The “Evers checklist” will be used to inform appraisal of the methodological quality of economic evaluations not applying economic modelling.⁵¹ The **quality of reporting** of the identified health economic studies will be assessed using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 24-item checklist.⁵² We will further include methods to incorporate an economic perspective into the GRADE assessment.⁵³ Two reviewers (KK, NN) will assess study quality. Disagreement will be solved in discussion with a third reviewer.

Synthesis of findings

We will summarize in a comparative way the available study data considering characteristics such as study design, location, setting, population and subpopulations (e.g., considering age,

vaccination status and others, if possible), different HPV screening strategies, the methodological quality, and outcomes including their results.

Assessment of the external validity/applicability to the Swiss setting

The applicability of the results to Switzerland will be evaluated considering the Criteria of the European network for Health Technology Assessment (EUnetHTA) HTA adaptation toolkit.⁵⁴

4.4.2 Cost-effectiveness analysis

We will consider a cost-effectiveness analysis using a decision-tree model to outline the existing evidence on cervical cancer screening considering different screening scenarios (e.g., including different screening test combinations, test types, types of sample collection, screening intervals [the final stratification will depend on the available data]) and characteristics of eligible women (e.g., different age groups, vaccination status) for Switzerland. Our primary analysis will focus on the following screening test comparisons (also illustrated in Figure 1):

Indextest

- HPV testing alone;
- HPV testing in combination with cytology (co-testing);
- HPV testing with cytology triage (conventional and LBC).

Comparison screening test

- Cytology-based testing without HPV triage;
- Cytology-based testing with HPV triage.

Reference standard

- Colposcopy with histological examination of tissue specimens;
- Colposcopy without histological examination of tissue specimens.

Cost-effectiveness analyses will combine the results of the effectiveness of different HPV screening strategies (as modelled using the decision-tree approach) in comparison to cytology-based testing with their corresponding costs (i.e., outcomes addressed in the clinical review [see section 4.2.5] will be contrasted with the costs of the respective screening strategy). Thereby, direct and indirect medical costs of the screening test strategies will be captured, including the costs of laboratory and diagnostic tests as well as test-related medical service utilization. Indirect non-medical costs, such as opportunity costs of test attendance, will not be considered in this analysis. Considering the currently published HTAs (Table 1, section 3) and/or other cost

analyses studies,⁵⁵⁻⁶³ we believe that direct evidence addressing the predefined outcomes will be available.

Outcomes

- Costs per LY gained; (discounted) ICER expressed as cost per LY gained;
- Costs per QALY gained; (discounted) ICER, expressed as cost per QALY gained;
- Costs per prevented case of cervical cancer-related death;
- Costs per prevented cervical cancer case (invasive cancer);
- Costs per prevented case of progressive precursor lesion;
- Costs per false-positive screening results (i.e., overdiagnosis and/or unnecessary treatment);
- Costs per false-negative screening results;
- Costs per colposcopy (with and without histological examination) referral;
- Costs per treatment;
- Costs related to psychological harms, adverse pregnancy outcomes, adverse treatment effects and any other reported harm;
- Incremental harm-benefit ratios (IHBR) expressed in units of additional harm per additional benefit;
- Overall strategy costs and subcategories of (overall) costs including their "drivers".

Outcomes will be stratified after different HPV screening strategies (e.g., considering different tests and test combinations, methods of sampling collection and screening intervals), the women's age and the vaccination status. Of note, the outcomes that will be considered in the cost-effectiveness analysis will follow an iterative process also depending on the data availability.

Perspective

The costs of the different screening strategies will be assessed from the healthcare providers' (payers') perspective.

4.4.3 Budget impact analysis

The aim of the BIA will be to investigate the overall costs of different HPV screening strategies (see above) in comparison to cytology-based screening in Switzerland. Age-dependent incidence rates (both for precursor lesions and invasive cervical cancer) will be combined with Swiss population statistics to outline the total costs of different primary HPV screening strategies in comparison to cytology-based screening.⁶⁴ Data will be assessed from the healthcare providers' (payers') perspective.

4.4.4 Sensitivity analysis

The base-case analysis will be conducted using the criteria defined and identified in the clinical review (see section 4.2). Analyses to address parameter uncertainty will include varying sets of related inputs including patient compliance or the HPV vaccination status.

Acknowledgement

We thank González A.I., MD, PhD (from the Institute of General Medicine, Johann Wolfgang Goethe University, Frankfurt, Germany) for her valuable input on the scoping report section addressing health-related preferences of women eligible for HPV screening.

5. References

1. Ferlay J, Ervik M, Lam F, et al. Global Cancer Observatory: Cancer Today. Secondary Global Cancer Observatory: Cancer Today. 2018. <https://gco.iarc.fr/today>.
2. World Health Organization, Department of Information, Evidence and Research. Mortality database. <http://www-dep.iarc.fr/WHODb/WHODb.htm> (Accessed 2019/08/20).
3. McCredie MR, Sharples KJ, Paul C, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol* 2008;9(5):425-34.
4. Crosbie EJ, Einstein MH, Franceschi S, et al. Human papillomavirus and cervical cancer. *Lancet* 2013;382(9895):889-99.
5. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999;189(1):12-9.
6. Melnikow J, Henderson JT, Burda BU, et al. Screening for Cervical Cancer With High-Risk Human Papillomavirus Testing: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA* 2018;320(7):687-705.
7. Évaluation de la recherche des papillomavirus humains (HPV) en dépistage primaire des lésions précancéreuses et cancéreuses du col de l'utérus et de la place du double immuno-marquage p16/Ki67. Saint-Denis La Plaine: Haute Autorité de Santé, 2019.
8. Gilham C, Sargent A, Kitchener HC, et al. HPV testing compared with routine cytology in cervical screening: long-term follow-up of ARTISTIC RCT. *Health Technol Assess* 2019;23(28):1-44.
9. Health Information and Quality Authority. Safer Better Care. Health technology assessment of human papillomavirus testing as the primary screening method for prevention of cervical cancer. https://www.hiqa.ie/sites/default/files/2017-05/HTA%20of%20cervical%20screening_plain%20english%20summary290517.pdf (Accessed 2019/08/25).
10. Arbyn M, Haelens A, Desomer A, et al. Cervical cancer screening program and Human Papillomavirus (HPV) testing, part II: update on HPV primary screening. *Health Technology Assessment (HTA), Reports 238Cs D/2015/10273/16 Brussels: Belgian Health Care Knowledge Centre*, 2015.
11. Nutzenbewertung eines HPV-Tests im Primärscreening des Zervixkarzinoms – Aktualisierung. S13-03 ed. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2014.
12. Koliopoulos G, Nyaga VN, Santesso N, et al. Cytology versus HPV testing for cervical cancer screening in the general population. *Cochrane Database Syst Rev* 2017;8:Cd008587.
13. Chao YS, Clark M, Carson E, et al. CADTH Optimal Use Reports. HPV Testing for Primary Cervical Cancer Screening: A Health Technology Assessment. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health. Canadian Agency for Drugs and Technologies in Health, 2019.
14. Higgins JPT, Savović J, Page MJ, et al. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook (Accessed 2019/10/01).
15. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155(8):529-36.
16. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
17. Wells GA, Shea BJ, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Secondary The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses 2014. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (Accessed 2019/08/10).

18. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med* 2001;20(3 Suppl):21-35.
19. Arbyn M, Smith SB, Temin S, et al. Detecting cervical precancer and reaching underscreened women by using HPV testing on self samples: updated meta-analyses. *BMJ* 2018;363:k4823.
20. McGowan J, Sampson M, Salzwedel DM, et al. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol* 2016;75:40-6.
21. Rebolj M, Rimmer J, Denton K, et al. Primary cervical screening with high risk human papillomavirus testing: observational study. *BMJ* 2019;364:l240.
22. Cook DA, Mei W, Smith LW, et al. Comparison of the Roche cobas(R) 4800 and Digene Hybrid Capture(R) 2 HPV tests for primary cervical cancer screening in the HPV FOCAL trial. *BMC Cancer* 2015;15:968.
23. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst* 2008;100(7):492-501.
24. Ogilvie GS, Krajden M, van Niekerk D, et al. HPV for cervical cancer screening (HPV FOCAL): Complete Round 1 results of a randomized trial comparing HPV-based primary screening to liquid-based cytology for cervical cancer. *Int J Cancer* 2017;140(2):440-48.
25. Zorzi M, Frayle H, Rizzi M, et al. A 3-year interval is too short for re-screening women testing negative for human papillomavirus: a population-based cohort study. *BJOG* 2017;124(10):1585-93.
26. Kitchener HC, Almonte M, Thomson C, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol* 2009;10(7):672-82.
27. El Sherif R, Pluye P, Gore G, et al. Performance of a mixed filter to identify relevant studies for mixed studies reviews. *J Med Libr Assoc* 2016;104(1):47-51.
28. NHS EED filter. Centre for Review and Dissemination, University of York. <https://www.crd.york.ac.uk/CRDWeb/Homepage.asp#nhseedmedline> (Accessed 2019/08/10).
29. Meijer CJ, Berkhof J, Castle PE, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *Int J Cancer* 2009;124(3):516-20.
30. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002;287(16):2114-9.
31. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
32. Sterne JAC, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.
33. Sterne JAC, Hernan MA, McAleenan A, et al. Chapter 25: Assessing risk of bias in a non-randomized study. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook (Accessed 2019/10/01).
34. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924-6.
35. Langer G, Meerpohl JJ, Perleth M, et al. GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. *Z Evid Fortbild Qual Gesundheitswes* 2012;106(5):357-68.
36. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64(4):383-94.
37. Guyatt GH, Oxman AD, Schunemann HJ, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol* 2011;64(4):380-2.
38. Reeves BC, Deeks JJ, Higgins JPT, et al. Chapter 24: Including non-randomized studies on intervention effects. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook (Accessed 2019/10/01).

39. Senn S. Trying to be precise about vagueness. *Stat Med* 2007;26(7):1417-30.
40. Guyatt GH, Thorlund K, Oxman AD, et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. *J Clin Epidemiol* 2013;66(2):173-83.
41. Deeks JJ, Higgins JPT, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook (Accessed 2019/10/01).
42. Higgins JPT, Eldridge S, Li T. Chapter 23: Including variants on randomized trials. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook (Accessed 2019/10/01).
43. Ebrahim S, Akl EA, Mustafa RA, et al. Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers. *J Clin Epidemiol* 2013;66(9):1014-21.e1.
44. Ebrahim S, Johnston BC, Akl EA, et al. Addressing continuous data measured with different instruments for participants excluded from trial analysis: a guide for systematic reviewers. *J Clin Epidemiol* 2014;67(5):560-70.
45. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. *Annu Rev Public Health* 2014;35:29-45.
46. Pearson A, White H, Bath-Hextall F, et al. A mixed-methods approach to systematic reviews. *Int J Evid Based Healthc* 2015;13(3):121-31.
47. Hong QN, Pluye P, Bujold M, et al. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Syst Rev* 2017;6(1):61.
48. Shemilt I, Mugford M, Byford S, et al. Chapter 15: Incorporating economics evidence. In: Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011] ed: The Cochrane Collaboration, 2011.
49. Philips Z, Bojke L, Sculpher M, et al. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* 2006;24(4):355-71.
50. Cooper N, Coyle D, Abrams K, et al. Use of evidence in decision models: an appraisal of health technology assessments in the UK since 1997. *J Health Serv Res Policy* 2005;10(4):245-50.
51. Evers S, Goossens M, de Vet H, et al. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care* 2005;21(2):240-5.
52. Husereau D, Drummond M, Petrou S, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *Value Health* 2013;16(2):e1-5.
53. Brunetti M, Shemilt I, Pregno S, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol* 2013;66(2):140-50.
54. European Network for Health Technology Assessment. EUnetHTA-HTA adaptation toolkit: work package 5; version 5 [online]. 2011: <https://www.eunethta.eu/> (Accessed 2019/08/20).
55. Sroczyński G, Siebert U. Decision Analysis for the Evaluation of Benefits, Harms and Cost-effectiveness of Different Cervical Cancer Screening Strategies to Inform the S3 Clinical Guideline "Prevention of Cervical Cancer" in the Context of the German Health Care System. University for Health Sciences, Medical Informatics and Technology. Austria, 2016.
56. Szucs TD, LARGERON N, Dedes KJ, et al. Cost-effectiveness analysis of adding a quadrivalent HPV vaccine to the cervical cancer screening programme in Switzerland. *Curr Med Res Opin* 2008;24(5):1473-83.
57. Vassilakos P, Poncet A, Catarino R, et al. Cost-effectiveness evaluation of HPV self-testing offered to non-attendees in cervical cancer screening in Switzerland. *Gynecol Oncol* 2019;153(1):92-99.
58. Vijayaraghavan A, Efrusy MB, Mayrand MH, et al. Cost-effectiveness of high-risk human papillomavirus testing for cervical cancer screening in Quebec, Canada. *Can J Public Health* 2010;101(3):220-5.

59. Kim JJ, Campos NG, Sy S, et al. Inefficiencies and High-Value Improvements in U.S. Cervical Cancer Screening Practice: A Cost-Effectiveness Analysis. *Ann Intern Med* 2015;163(8):589-97.
60. Bains I, Choi YH, Soldan K, et al. Clinical impact and cost-effectiveness of primary cytology versus human papillomavirus testing for cervical cancer screening in England. *Int J Gynecol Cancer* 2019.
61. Hylén H, Thrane H, Pedersen K, et al. The healthcare costs of treating human papillomavirus-related cancers in Norway. *BMC Cancer* 2019;19(1):426.
62. Lee H, Hur S, Jang H, et al. Cost-Utility of a Two-Dose Human Papillomavirus Vaccination Programme Added to Cervical Cancer Screening Compared with Cervical Cancer Screening Alone in Korea. *Asian Pac J Cancer Prev* 2019;20(2):425-35.
63. Termrungruanglert W, Khemapech N, Tantitamit T, et al. Cost effectiveness analysis of HPV primary screening and dual stain cytology triage compared with cervical cytology. *J Gynecol Oncol* 2019;30(2):e17.
64. Ochs K, Meili G, Diebold J, et al. Incidence Trends of Cervical Cancer and Its Precancerous Lesions in Women of Central Switzerland from 2000 until 2014. *Front Med (Lausanne)* 2018;5:58.

Appendix A. Pilot search strategy.

Pilot search strategy for the clinical review in Medline using PubMed for cervical cancer screening with human papillomavirus testing performed on the 06th of August 2019.

Cervical Cancer	(Uterine Cervical Neoplasms [mh] OR Uterine Cervical Dysplasia [mh] OR Cervical Intraepithelial Neoplasia [mh] OR ((cervix [tw] OR cervical [tw] OR cervico* [tw]) AND (cancer* [tw] OR carcinoma OR adenocarcinoma OR neoplas* [tw] OR dysplas* [tw] OR dyskaryos* [tw] OR squamous [tw] OR CIN [tw] OR CINII* [tw] OR CIN2* [tw] OR CINIII* [tw] OR CIN3* [tw] OR SIL [tw] OR HSIL [tw] OR H-SIL [tw] OR LSIL [tw] OR L-SIL [tw] OR ASCUS [tw] OR AS-CUS [tw])))
	AND
HPV	(papillomaviridae [mh] OR alphapapillomavirus [mh] OR DNA, viral [mh] OR Papillomavirus Infections [mh] OR Tumor Virus Infections [mh] OR Cervix Uteri/virology [mh] OR HPV [tw] OR “human papillomavirus” [tw] OR papillomaviridae [tw] OR PCR OR “hybrid capture*” [tw] OR HC2 [tw] OR HCII [tw] OR “HC 2” [tw] OR “HC II” [tw] OR ((viral [tw] OR virolog* [tw]) AND DNA [tw]))
	AND
Cytology	(Vaginal smears [mh] OR Cytodiagnosis [mh] OR Cell Transformation, Viral [mh] OR Cytopathogenic Effect, Viral [mh] OR ((pap [tw] OR papanicolaou [tw] OR vagina* [tw] OR cervical [tw] OR cervix [tw] OR cervico* [tw] OR cytolog* [tw]) AND (smear* OR test [tw] OR tests [tw] OR testing [tw] OR tested [tw] OR swab* OR scrap*)))
	AND
Screening	(Screening [tw] OR screen* OR Early Detection of Cancer [mh] OR Mass Screening [mh])

mh: Mesh term; tw: text word.