FACULTY OF MEDICINE AND HEALTH SCIENCES

Evidence regarding Human Papillomavirus Testing in Secondary Prevention of Cervical Cancer

Marc Arbyn

Scientific Institute of Public Health, Unit of Cancer Epidemiology/Belgian Cancer Centre J. Wytsmanstreet 14, B1050 Brussels, Belgium

Promoters:
Prof. Dr. M. Temmerman[1]
Prof. Dr. S. Weyers[1]
Prof. Dr. J. Bogers[1,2]

[1] Department of Obstetrics and Gynaecology, Ghent University;
[2] Applied Molecular Biology Research Group (AMBIOR), Laboratory for Cell and Tissue Research, University of Antwerp

Doctoral thesis submitted to the Faculty of Medicine and Health Sciences, Ghent University, Belgium

GHENT UNIVERSITY

ICRH

ISBN-NUMMER : 9789078128472
EAN : 9789078128472

FACULTY OF MEDICINE AND HEALTH SCIENCES
GHENT UNIVERSITY
BELGIUM

Evidence regarding Human Papillomavirus Testing in Secondary Prevention of Cervical Cancer

Marc Arbyn

Scientific Institute of Public Health, Unit of Cancer Epidemiology, Belgian Cancer Centre, J. Wytsmanstreet 14, B1050 Brussels, Belgium

Promoters:
Prof. Dr. M. Temmerman[1]
Prof. Dr. S. Weyers[1]
Prof. Dr. J. Bogers[1,2]

[1] Department of Obstetrics and Gynaecology, Ghent University; [2] Applied Molecular Biology Research Group (AMBIOR), Laboratory for Cell and Tissue Research, University of Antwerp

Doctoral thesis submitted to the Faculty of Medicine and Health Sciences, Ghent University, Belgium, 13th of November, 2017
ISBN-NUMMER : 9789078128472
EAN : 9789078128472

1. Table of contents

1. Table of contents 3
2. List of abbreviations 7
3. Executive summary 9
4. Samenvatting 11
5. Introduction 14
   5.1. Human papillomaviruses 14
   5.2. Association between human papillomavirus (HPV) i ènfection and cervical (pre-) cancer and other HPV-related cancers 14
   5.3. Cervical cancer screening using the Pap smear 17
   5.4. Applications of HPV testing in clinical practice 20
   5.5. Background regarding HPV testing to triage women with minor cervical cytological lesions 20
   5.6. Background regarding HPV testing as a test of cure after treatment for cervical pre-cancer 21
   5.7. Background regarding HPV testing in primary cervical cancer screening 22
6. Objectives 23
   6.1. Primary objective 23
   6.2 Secondary objectives 23
   6.2.1. Triage of women with minor abnormal cervical cytology 23
   6.2.2. Follow-up of women treated for cervical pre-cancer 23
   6.2.3. Primary cervical cancer screening 24
   6.2.4. HPV testing on self-samples 24
7. Methods 25
   7.1. Definitions and terminology 25
   7.2. Meta-analysis: purpose and methods 25
   7.3. Literature retrieval 25
   7.4. Eligibility of studies 26
   7.4.1. Triage of women with minor abnormal cervical cytology 26
   7.4.2. Follow-up of women treated for cervical pre-cancer 26
   7.4.3. Primary cervical cancer screening 27
   7.4.4. HPV testing on self-samples 27
   7.5. Targeted outcomes 27
   7.6. Reference or gold standard 28
   7.7. Data extraction and management 28
   7.8. Outcome measures 28
   7.9. Methodological quality 28
   7.10. Statistical analysis 29
   7.10.1. Absolute accuracy 29
   7.10.2. Relative accuracy 29
   7.10.3. Pooling of proportions or ratios 29
   7.10.4. Assessment of heterogeneity and sources of heterogeneity 29
   7.11. Presentation of the results 30
8. Triage of women with minor abnormal cervical cytology 31
   8.1. Retrieved studies 31

8.2. Triage of ASC-US 32
8.2.1. Hybrid Capture II (HC2) 32
8.2.2. Other HPV assays and biomarkers 34
8.3. Triage of LSIL 34
8.3.1. Hybrid Capture II 34
8.3.2. Other HPV assays and biomarkers 35
8.4. Requirements for a test usable in triage of minor cytological lesions 36
8.4.1. hrHPV testing in triage of ASC-US 36
8.4.2. hrHPV testing in triage of LSIL 36
8.5. Tables: Triage of women with minor abnormal cervical cytology 37
9. Follow-up of women treated for cervical pre-cancer 43
9.1. Retrieved studies 43
9.2. Treatment failure 43
9.3. Absolute accuracy of hrHPV-DNA testing to predict occurrence of CIN2+ after treatment for pre-cancer. 44
9.4. Absolute accuracy of cytology follow-up to predict residual/recurrent CIN2+. 46
9.5. Absolute accuracy of combined hrHPV-DNA and cytology testing to predict occurrence of CIN2+ after treatment for pre-cancer 47
9.6. Relative accuracy of hrHPV-DNA testing versus cytology to predict occurrence of CIN2+ after treatment for pre-cancer. 48
9.7. Relative accuracy of combined hrHPV-DNA and cytology testing versus hrHPV-DNA testing alone to predict occurrence of CIN2+ after treatment for pre-cancer. 49
9.8. Age as a risk factor for occurrence of disease after treatment 49
9.9. Tables: Follow-up of women treated for cervical precancer 51
10. Primary cervical cancer screening 52
10.1. Clinical accuracy of HPV tests to detect CIN2+ and CIN3+ 52
10.1.1. Retrieved studies 52
10.1.2. Absolute cross-sectional accuracy of HPV testing 52
10.1.3. Relative cross-sectional accuracy 54
10.1.4. Cross-sectional outcomes of randomised trials comparing cytology- with HPV based screening 55
10.2. Efficacy and effectiveness of HPV vs cytology in primary screening for cervical cancer 57
10.2.1. Retrieved studies 57
10.2.2. Reduction in the cumulative incidence of CIN3+ and of cervical cancer 57
10.3. Triage of HPV positive women 60
10.3.1. Absolute accuracy of triage algorithms 60
10.3.2. Relative accuracy of cytology and/or hrHPV-DNA based triage algorithms versus reflex cytology (ASC-US+) 63
10.3.3. $p16^{INK4a}$ immuno-cytochemistry 64
10.4. Screening interval 65
10.5. Target age group for HPV-based screening 68
10.5.1. Principles 68
10.5.2. Age-specific prevalence of HPV infection 68

10.5.3. Variation in test accuracy of cervical cytology and hrHPV DNA testing by age. 69
10.5.4. Influence of age on screening efficacy 70
10.5.5. Age to stop screening 70
10.5.6. Age-specific adverse effects related to treatment of screen-detected lesions 70
10.6. Validation of HPV assays for primary cervical cancer screening 70
10.7. Tables: Primary screening 73
11. HPV testing on self-samples 83
11.1 Retrieved studies 83
11.2. Assessment of study quality 83
11.3. Absolute sensitivity and specificity, by clinical setting 84
11.4. Relative sensitivity and specificity of HPV testing in self-samples versus testing on clinician-samples 85
11.5. Effects of the HPV assay and of the self-sampling device 85
11.6 Tables: HPV testing on self-samples 89
12. Discussion 91
12.1. Triage of women with ASC-US 91
12.2. Triage of women with LSIL 92
12.3. Utility of HPV testing in a given setting derived from PPP plots 93
12.4. Accuracy of HPV-DNA testing and cytology to predict treatment failure 98
12.5. Accuracy of HPV testing in cervical cancer screening 99
12.6. Effectiveness of HPV testing in cervical cancer screening 100
12.6.1. Existence of HPV-negative cervical cancers: a reason to prefer co-testing? 100
12.6.2. Longitudinal criteria to demonstrate effectiveness of other screening tests which are not HPV DNA tests 103
12.6.3.Cervical cancer screening of vaccinated cohorts 106
12.6.4. Cost-effectiveness of HPV-based screening 106
12.7. HPV testing on self-samples 107
12.8. Limitations of meta-analyses 108
13. Conclusions and propositions for further research 110
13.1. Conclusions 110
13.1.1. Triage of women with minor abnormal cervical cytology 110
13.1.2. hrHPV testing as a test of cure after treatment of cervical pre-cancer 111
13.1.3. hrHPV testing in primary cervical cancer screening 111
13.1.4. HPV testing on self-samples 112
13.2. Propositions for further research 112
13.2.1. Triage of women with minor abnormal cervical cytology 112
13.2.2. Follow-up of women treated for cervical pre-cancer 113
13.2.3. Primary cervical cancer screening 113
13.2.4. HPV testing on self-samples 114
13.3. Internationalisation of health technology assessment 115
14. Acknowledgements 116
15. Annexes 117
15.1. Terminology used for reporting of cervical cytological or histological findings 117

15.2. Virological test characteristics of HPV assays evaluated in the PhD project 118
15.3. Search strings 124
15.4. Studies included in the meta-analysis of the accuracy of HPV testing in triage of women with minor cytological abnormalities 127
15.4.1. Triage of ASC-US 127
15.4.2. Triage of LSIL 131
16. Biography and bibliography 134
17. Attached publications 137
18. References 202

2. List of abbreviations

AGC	atypical glandular cells
AGUS	atypical glandular cells of undetermined significance
AIS	adenocarcinoma in situ
ASC-H	atypical squamous cells cannot exclude HSIL
ASC-US	atypical squamous cells of undetermined significance
BSCC	British Society of Clinical Cytology
CGIN	cervical glandular intraepithelial neoplasia
CI	confidence interval
CIB	confidence interval bound
CIN	cervical intraepithelial neoplasia
CIS	carcinoma in situ
cNPV	complement of the negative predictive value (=1-NPV)
CO	cut-off
CT	cycle threshold
DOR	diagnostic odds ratio
DR	detection rate
DTA	diagnostic test accuracy
EIA	enzyme immune-assay
ESS	effective sample size
HC	Hybrid Capture
HC2	Hybrid Capture II assay
HPV	human papillomavirus
hr	high-risk
hsROC	hierarchical summary receiver-operating-characteristics
ISP	Institut scientifique de Santé publique
LAST	Lower Anogenital Squamous Terminology Standardization
LBC	liquid-based cytology
LSIL	low-grade squamous epithelial lesions
MALDI-TOF	matrix-assisted laser desorption-ionization time-of-flight
NILM	Negative for epithelial abnormality
NLR	negative likelihood ratio NPV: negative predictive value
PCR	polymerase chain reaction
PLR	positive likelihood ratio
PPP plot	pretest posttest probability plot
PPV	positive predictive value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	randomised clinical trial
RR	relative risk
RLU	relative light units
RT-PCR	real-time polymerase chain reaction
SSSCCA	superficially invasive squamous cell carcinoma

sROC	summary receiver-operating-characteristics
TBS	The Bethesda System
VALGENT	VALidation of HPV GENotyping Tests
VALHUDES	VALidation of Human papillomavirus assays and collection DEvices for HPV testing on Self-samples and urine samples
WIV	Wetenschappelijk Instituut Volksgezondheid

3. Executive summary

Introduction
The large majority of cervical cancers are caused by a persistent infection of the cervical epithelium with human papillomavirus (HPV) types, which are transmitted mainly through sexual contact. The carcinogenic potential has been demonstrated for a dozen of HPV types, which therefor are termed as high-risk (hr).
The PhD thesis assesses the evidence for four clinical applications of hrHPV testing in the framework of the secondary prevention of cervical cancer.

Methods
The four clinical applications of hrHPV testing which we evaluated subsequently as considered historically in the scientific literature are: (1) triage of women with minor abnormal cervical cytology, (2) follow-up of women treated for cervical pre-cancer, (3) primary cervical cancer screening, and (4) HPV testing on self-samples collected by the woman her-self.
We conducted systematic reviews and meta-analyses to assess a series a clinical questions, which in essence compared the performance of a conventional and a new method of intervention or testing. Two main designs were applied: (1) evaluation of the absolute or relative diagnostic test accuracy (DTA) and (2) evaluation of the efficacy or effectiveness of an intervention. DTA reviews were performed for all the four clinical applications, where we tried to derive the average sensitivity and specificity of hrHPV tests, as well as the relative sensitivity and specificity of hrHPV assays versus other assays to detect cervical pre-cancer. Cervical pre-cancer is defined as histologically confirmed presence of cervical intraepithelial neoplasia of grade II or III or adenocarcinoma in situ of the cervix uteri. Questions related to the impact of hrHPV-based versus cytology-based screening for cervical cancer were addressed by pooling randomised trials and screening cohort studies.
Established procedures, recommended by the Cochrane Collaboration and health-technology assessment guidelines, were followed to retrieve the key data from selected relevant references from the scientific literature and to synthesize findings from many studies into overall measures. The quality of selected studies and variability of results were examined and accounted for as far as possible.

Results
hrHPV testing is more accurate (more sensitive and similarly specific) than repeat cytology to find underlying cervical pre-cancer in women with a cytological finding of atypical squamous cells of undetermined significance (ASC-US). Triage of women with low-grade squamous intraepithelial lesions (LSIL) with a hrHPV DNA test is sensitive but poorly specific. The utility of hrHPV testing in women with LSIL is limited, since the posttest risk of pre-cancer in case of LSIL is hardly higher than the pretest risk. Triage of ASC-US or LSIL with a RNA test targeting all high-risk HPV types is as sensitive but more specific than the Hybrid Capture II assay (the most used hrHPV DNA test). Cyto-immuno-chemistry for p16 or p16/ki67 is substantially more specific but slightly less sensitive to find underlying pre-cancer in women with ASC-US or LSIL. Isolated

HPV16/18 genotyping is insufficiently sensitive in triage of women with minor abnormal cytology but distinction between HPV16/18 positive women, other hrHPV positive womenstratifies risk allowing for specific follow-up recommendations for hrHPV positive women.
Pretest-Posttest-Probability plots are useful to demonstrate the clinical utility of available triage options.

Women treated for cervical pre-cancer are at significantly increased risk of recurrent cervical (pre-) cancer compared to the general population and therefore require careful follow-up. hrHPV testing after approximately halve a year post-treatment is more sensitive and almost as specific as repeat cytology to predict residual of recurrent CIN2+. Adding cytology (co-testing) increases the sensitivity but decreases the specificity of the prediction of treatment failure. The risk of CIN3 in case of a double negative co-test 6 and 18 months after treatment is as low as for women with a negative cytological screening test.

Primary cervical cancer screening with hrHPV tests is more sensitive but less specific for detecting underlying cervical pre-cancer than cytology. Moreover, randomised trials have consistently shown that the subsequent risk of CIN3 and of invasive cervical cancer is lower five years after a negative hrHPV DNA test compared to three years after a negative cytology result. Prevalence of transient hrHPV infection is high in women younger than 30 years and no increased efficacy of hrHPV screening in this age group is demonstrated. Improved protection against cervical cancer by hrHPV DNA screening using HC2 or GP5+/6+ PCR-EIA has been directly observed in randomised trials. It has been accepted that other hrHPV DNA tests, which demonstrate good reproducibility and non-inferior cross-sectional sensitivity and specificity for CIN2+ compared to HC2 or GP5+/6+ PCR-EIA, can also be applied in primary cervical cancer screening. The following hrHPV DNA assays fulfil these equivalency criteria; cobas-4800, Abbott RealTime High Risk HPV, BD Onclarity, PapilloCheck, and the HPV-Risk assay. The APTIMA HPV assay is highly reproducible, as sensitive and more specific than the standard comparator assays, but –since it targets mRNA of hrHPV - requires additional demonstration of longitudinal safety over 5-year or longer intervals. hrHPV positive women require triage before referral for colposcopy. Reflex cytology followed, six to twelve months later by a second cytology test or hrHPV retest when reflex-cytology is negative, completed with HPV16/18 genotyping or not, show good sensitivity with moderate to low specificity for detecting cervical pre-cancer among hrHPV positive women. P16/Ki67 chemistry and certain methylation markers may provide promising alternative triage options.

hrHPV DNA testing on self-samples is as accurate as on clinician-taken samples under the condition that a validated PCR-based assay is used.

Discussion and conclusions
hrHPV testing can be recommended for triage of women with ASC-US and as test of cure after treatment for cervical pre-cancer.

Considering available evidence, European guidelines propose cervical cancer screening in the Member States with a clinically validated hrHPV DNA assay, followed by an appropriate triage algorithm to manage hrHPV positive women. Validated PCR platforms detecting hrHPV DNA can be used also on self-samples, which may generate opportunities to reach women who do not participate in regular screening.
The existence of hrHPV negative cancers that may be missed by HPV-only screening and the question whether co-testing with hrHPV and cytology may avoid these cases is the object of intensive debate. Careful cost-effectiveness analyses together with new meta-analytical tools for pooling of cumulative incidence data according to the baseline co-test status, may open pathways to address these questions. Pooling of cumulative incidence curves may also provide a key for defining the longitudinal criteria for new screening tests that target other molecules than hrHPV DNA.

4. Samenvatting

Inleiding
De grote meerderheid van baarmoederhalskankers zijn te wijten aan persistente infectie van het cervixepitheel met types van het humaan papillomavirus (HPV), die vooral door seksueel contact worden overgedragen. Een twaalftal HPV types zijn potentieel kankerverwekkend en worden daarom beschouwd als hoog-risico (hr) types.
Deze PhD thesis heeft de bedoeling om de evidentie voor vier mogelijke klinische toepassingen van hrHPV testing te evalueren in het kader van secondaire preventie van baarmoederhalskanker.

Methodes
De vier toepassingen voor hrHPV testing die we geëvalueerd hebben in de chronologische volgorde dat ze werden overwogen in de wetenschappelijke literatuur zijn: (1) triage van vrouwen met een geringe cytologische afwijkingen van het cervixepitheel, (2) opvolging van vrouwen behandeld voor cervix-prekanker, (3) primaire screening naar baarmoederhalskanker, en (4) HPV testing op een vaginaal staal afgenomen door de vrouw zelf.
We hebben verscheidene systematische reviews en meta-analyses uitgevoerd om een serie van klinische vragen te beantwoorden die de performantie van conventionele en nieuwe testen of interventies vergeleken. Twee studieontwerpen werden toegepast: (1) evaluatie van diagnostische testen en (2) evaluatie van de effectiviteit van interventies. Reviews van de diagnostische accuraatheid werden uitgevoerd voor al de vier klinische toepassingen, waar we geprobeerd hebben om de absolute gevoeligheid en specificiteit van hrHPV testen te evalueren, alsook van de relatieve gevoeligheid en specificiteit in vergelijking met andere testen. De uitkomst was detectie van precursoren van baarmoederhalskanker of invasieve baarmoederhalskanker. Met precursoren bedoelen we histologisch geconfirmeerde aanwezigheid van cervicale intraepitheliale neoplasie (CIN) van graad II of III of adenocarcinoma in situ. Vragen over de effectiviteit van screening met hrHPV testing versus cytologische screening werden aangepakt door data te beschouwen van gerandomiseerde steekproeven en cohortestudies.

Standaard procedures, aanbevolen door de Cochrane Collaboratie en in richtlijnen voor evaluatie van nieuwe technologieën werden opgevolgd voor het vinden van data uit geselecteerde relevante wetenschappelijke artikels en om de evidentie, geëxtraheerd uit een veelheid van studies, te synthetiseren. De kwaliteit van gevonden studies en de variatie van de resultaten werden zorgvuldig onderzocht en in rekening gebracht voor zover als mogelijk.

Resultaten
hrHPV testing is adequater (meer gevoelig en even specifiek) dan herhalingscytologie om onderliggende cervix pre-kanker te detecteren in vrouwen met ASC-US[a]. Triage van vrouwen met LSIL[b] is gevoelig maar heel weinig specifiek. Het nut van hrHPV testing bij vrouwen met LSIL is beperkt, vermits het posttest risico van cervicale prekanker nauwelijks hoger is dan the pretest risico.
Triage van ASC-US of LSIL met een RNA test, die alle hrHPV types detecteert, is even gevoelig maar specifieker om cervix prekanker op te sporen dan de Hybrid Capture II test (de meest gebruikte hrHPV DNA test). De geïsoleerde bepaling van de twee meest oncogene HPV types (HPV16/18) is te weinig gevoelig voor triage van vrouwen met geringe cytologische afwijkingen. Daarentegen, kan het onderscheid tussen infectie met HPV16/18 of infectie met ander hrHPV types nuttig zijn voor een specifiek follow-up beleid van hrHPV-positieve vrouwen.
De Pretest-Posttest-Risico grafiek laat toe om het nut van diverse triage opties te visualiseren.

Vrouwen behandeld voor een hooggradig CIN letsel hebben een aanzienlijk verhoogd risico om nadien cervix (pre-) kanker te ontwikkelen en daarom is zorgvuldige opvolging nodig. hrHPV testing ongeveer zes maanden na behandeling is gevoeliger en bijna even specifiek als herhalingscytologie om een recidief van CIN2+ te voorspellen. Tegelijk ook een cytologie test toevoegen verhoogt de sensitiviteit maar verlaagt de specificiteit om therapeutisch falen aan te tonen. Het risico op een CIN3 letsel, volgend op een negatieve co-test 6 en 18 maanden na de behandeling, is even laag als voor vrouwen met een negatieve cytologische screeningstest.

Primaire screening naar baarmoederhalskanker met hrHPV testen is gevoeliger maar minder specifiek om onderliggende cervix prekanker op te sporen dan cytologie. Bovendien hebben gerandomiseerde studies herhaaldelijk aangetoond dat de kans op ontwikkeling van CIN3 of invasieve kanker lager is vijf jaar na een negatieve hrHPV test dan drie jaar na een negatieve cytologie. De prevalentie van voorbijgaande hrHPV infectie is hoog bij vrouwen jonger dan 30 jaar en de doeltreffendheid van hrHPV screening in deze groep is niet aangetoond.
Verbeterde bescherming tegen cervix uteri kanker door screening met hrHPV testing is waargenomen in gerandomiseerde studies waar de HC2 en de GP5+/6+ PCR-EIA werden gebruikt. Andere hrHPV DNA testen kunnen ook worden aanvaard voor baarmoederhals-kankerscreening indien een hoge reproduceerbaarheid en ook een sensitiviteit en specificiteit voor CIN2+ kan aangetoond worden die niet lager is dan HC2 of GP5+/6+

[a] ASC-US: atypische squameuze cellen van onbepaalde betekenis.
[b] LSIL: laaggradige squameuze intra-epitheliale laesie.

PCR-EIA. De volgende hrHPV DNA testen vervullen deze equivalentie criteria; cobas-4800, Abbott RealTime High Risk HPV, BD Onclarity, PapilloCheck, en de HPV-Risk assay.

De APTIMA HPV test is heel reproduceerbaar, even gevoelig en meer specifiek dan de HC2 of GP5+/6+ PCR-EIA, maar vereist bijkomend bewijs van longitudinale veiligheid over een periode van 5 jaar of langer om deze test te kunnen aanbevelen voor primaire screening. hrHPV positieve vrouwen moeten eerst een triage ondergaan om uit te maken wie verder colposcopisch onderzoek nodig heeft. Reflex-cytologie gevolgd zes tot twaalf maanden later door een tweede cytologie test ofwel een hrHPV test zes tot twaalf maanden later als reflex-cytologie negatief was, met of zonder HPV16/18 genotypering, tonen een goede sensitiviteit maar een matige tot zwakke specificiteit. P16/Ki67 immunocytochemistry en zekere methylatie merkers hebben belovende resultaten getoond voor triage van hrHPV positieve vrouwen.

hrHPV DNA testing op zelf-afgenomen vaginale stalen is even accuraat als op stalen afgenomen door een clinicus op voorwaarde dat een gevalideerde PCR test wordt gebruikt.

Discussie en besluiten
hrHPV testing kan worden aanbevolen voor triage van vrouwen met ASC-US en bij opvolging na behandeling van cervix prekanker.
Op basis van beschikbare evidentie, raden Europese richtlijnen aan om te screenen naar baarmoederhalskanker met een gevalideerde hrHPV DNA test, gevolgd door een adequate triage van hrHPV positieve vrouwen. Het gebruik van gevalideerde PCR testen op een zelf-afgenomen staal biedt de mogelijkheid om vrouwen te bereiken die niet deelnemen aan het screeningsprogramma.
Het bestaan van hrHPV negatieve kankers van de cervix uteri en de vraag of co-testing met hrHPV en cytologie deze kankers kunnen vermijden is het onderwerp van intensief debat en onderzoek. Zorgvuldige kosten-effectiviteitsstudies en nieuwe meta-analytische procedures om cumulatieve incidentie data te poolen in functie van de initiële co-test screening status, kunnen in de toekomst mogelijke antwoorden bieden op deze vragen. Pooling van cumulatieve incidentie data leveren mogelijks ook de sleutel voor de formulering van longitudinale criteria voor nieuwe screeningtesten die ander moleculen detecteren dan hrHPV DNA.

5. Introduction

5.1. Human papillomaviruses

Papillomaviruses (PV) are small, icosahedral DNA non-enveloped viruses , that consist of one single double stranded circular DNA molecule of approximately 8000 base-pairs, contained within a protein capsid[1]. The capsid is composed of two structural proteins, which both are encoded by the viral genome: L1 and L2. The viral genome contains eight open reading frames which are organised in three parts: 1) the early region (E), that encodes for six to seven E proteins (E1-E7) which are necessary for viral replication and may cause pathogenic effects in host cells, 2) the late region that encodes for the two structural capsid proteins, and 3) a long non-coding control region.

Papillomaviruses infect diverse vertebrates and are strictly host species specific[2]. PVs have evolved in linkage with their host and have shown very limited genetic change during their evolutionary history because of their typical location in epithelia which are less exposed to immunologic challenge[3]. Several hundreds of PV types have been identified that infect mucosal or cutaneous epithelia in humans only (HPV)[4]. PVs form a virus family on their own, the *Papillomaviridae,* and are grouped by genera, indicated by a Greek letter[5]. Genera are further subdivided into species and types. Taxonomic classification is based on phylogenetic homology and nucleotide sequence alignments [6]. Over hundred identified HPV types have been isolated from human tissues. They all belong to five genera (α, β,γ,μ,ν)[3,4].

5.2. Association between human papillomavirus (HPV) infection and cervical (pre-) cancer and other HPV-related cancers

In this PhD manuscript, we used the CIN terminology[7] to describe histological findings, whereas the SIL terminology of The Bethesda System (version 2001) was used to describe cytological lesions[8]. Pre-cancer was defined as histologically confirmed lesion, graded as CIN2 (cervical intraepithelial neoplasia of grade II), CIN3 (cervical intraepithelial neoplasia of grade II), possibly including also CGIN (cervical glandular and intraepithelial neoplasia) and AIS (adenocarcinoma in situ). A conversion table of different classification systems used for grading cytological and/or histological epithelial lesion can be found in annex 15.1.

The development of cervical cancer passes through a number of phases: (a) infection of the cervical epithelium with certain HPV types; (b) persistence of the HPV infection; (c) progression to pre-cancerous lesions (CIN) and (d) invasion. All phases can regress spontaneously, except the last one [1,9,10].

The precursors of cervical cancer arise mainly in the transformation zone (TZ) between the ectocervical multilayer squamous epithelium and the endo-cervical columnar epithelium[11,12] [11-14]. Recent histo-immunochemic research demonstrated that HPV-induced carcinogenesis originate from a discrete population of embryonic cells which are typically present at the squamo-cylindric junction in the uterine cervix (and probably also in the anal and oro-pharyngeal area)[15,16].

Recently, an IARC (International Agency for Research on Cancer) expert group reviewed the carcinogenicity of HPV and confirmed that for 12 alpha HPV types (HPV 16, 18, 31,

33, 35, 39, 45, 51, 52, 56, 58, and 59) sufficient evidence exists that they are causally linked with the development of cervical cancer [17] . Type HPV 68 is considered as probably carcinogenic (see Table 1)[18]. Some other HPV types may in rare occasions also cause cervical cancer (Figure 2)[19].
HPV type 16, in particular, has a high potential for malignant transformation of infected cervical cells[20]. The HPV types 16 and 18 jointly cause seven out of ten cervical cancers worldwide [21].
The next most important five types (HPV31, HPV33, HPV45, HPV52 and HPV58) together with HPV16/18 are causally linked with approximately 90% of cervical cancers[19]. Moreover, HP16 especially, is also linked with rarer types of cancer, such as cancer of the vulva and vagina in women, penis cancer in men and anal and oropharyngeal cancer in women and men[1,22].
In this manuscript, the term "high-risk" HPV types is used for the 13 oncogenic HPV types (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68) completed or not with HPV 66[c].

The carcinogenic effect of high-risk HPV types is driven by viral oncoproteins which interact with diverse molecular targets that deregulate the normal cell cycle[23].
The interactions of p53 and retinoblastoma protein (pRB) with the onco-proteins E6 and E7 are the most important ones[24]. E6 inhibition of p53 blocks apoptosis, whereas E7 inhibition of pRB abrogates cell cycle arrest. The expression of both oncoproteins is negatively controlled by E2. The insertion of viral DNA in to the human genome, which often occurs at the level of the E2 gene, may result in increased expression of the oncogenes E6 and E7, facilitating the carcinogenetic transformation of infected cells[25-28].
Integration may also interrupt other viral genes (L1 or E1) and by consequence, HPV may become non-detectable by tests targeting these genes [29]. Continuous E6 and E7 expression is believed to be necessary for the maintenance of the malignant phenotype of cervical cancer cells[28,30]. Subsequent to the altered transcription of E6/E7, a series of other molecular processes are triggered, resulting in genomic instability, epigenetic changes (hyper-methylation of certain viral and host genes; deregulated micro RNA expression), over-expression of certain cell-cycling regulating proteins and modified cytokine profiles. Unravelling of the carcinogenesis and bioinformatic processing of transcriptomic, proteomic and genomic profiles has revealed several biomarkers that may be used in triage and management of screen-positive women and may even identify new therapeutic targets [31-35] (see Figure 1).

The low-risk HPV types 6 and 11 cause approximately 90% of genital warts in women and men[36]. They occur in low-grade dysplastic cervical lesions, but are not associated with the development of cervical cancer[1].

[c] HPV66 was according to the 90th Volume of the IARC Monographs on the Evaluation of Carcinogenic Risks to Humans considered as carcinogenic (Lyon, 2007). However, in Volume 100, this type was included in the group 2B of possibly carcinogenic HPV viruses (Lyon, 2012).

Table 1. HPV types belonging to species α5, α6, α7, α9, and α11 according to the level of evidence of carcinogenicity for cervical cancer, expressed as colour shadings (Arbyn et al, *J Pathol* 2014[19], adapted from IARC Monograph, vol 100B[28]).

Species	Types							
α5	26	51	69	82				
α6	30	53	56	66				
α7	18	39	45	59	68	70	85	97
α9	16	31	33	35	52	58	67	
α11	34	73						

Group 1 carcinogens
Group 2A carcinogens
Group 2B carcinogens
Phylogenetic analogy with carcinogenetic types

The main route of HPV transmission is sexual contact. Infection usually occurs soon after the onset of sexual activity [37,38]. The prevalence of HPV infection in females generally peaks in late teenage or early twenties and declines thereafter [39]. HPV infection usually clears spontaneously, in particular in women younger than 30 years [40]. HPV infection can result in intraepithelial neoplastic cellular lesions (pre-cancers), which are identifiable by cytological examination and can be confirmed histologically. These lesions generally regress, but the probability of [1]regression decreases, and the likelihood of progression to cancer increases, with the duration of HPV infection and the severity of the lesion. From historical data, it has been estimated that CIN3 incurs a probability of progressing to invasive cancer of 12% to 30%; whereas for CIN2 this probability is substantially less[41,42].The progression of HPV infection to invasive cancer is thought to take a minimum of 10 years, but usually takes 15-30 years[43].

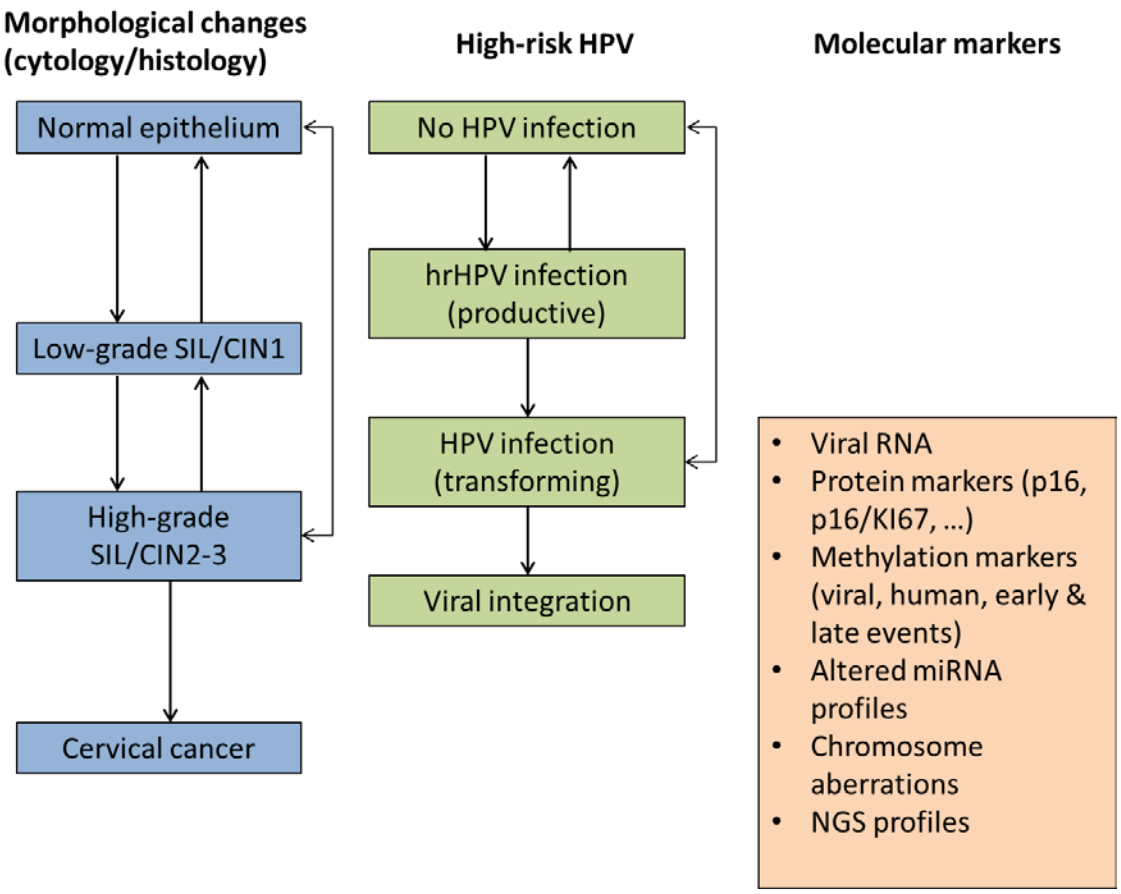


Figure 1. Diagram of the pathogenesis of cervical cancer, indicating tests that can be used in screening and triage of screen-positive women in order to detect cervical pre-cancerous lesions.

In many developed countries, the incidence of, and mortality from, squamous cervical cancer has dropped substantially over the last decades as a consequence of widespread screening [43-45]. For instance, in the USA and North and West-Europe, the cumulative risk of cervix uteri cancer up to age 74 years is 0.8% or lower [46,47]. However, approximately 54,000 and 11,000 cases are reported each year in Europe and the USA, respectively [2,48]. Moreover, the incidence of adenocarcinoma of the cervix is less affected by cytological screening [49,50]. In contrast to many other malignancies, cervical cancer primarily affects younger women, with the majority of cases appearing between the ages of 35 and 50 years [51,52]. In Europe, approximately 40% of women with cervical cancer die from the disease within five years of diagnosis [53].

5.3. Cervical cancer screening using the Pap smear

Among all malignant tumours, cervical cancer is the one which can be most effectively controlled by screening. Detection of cytological abnormalities by microscopic examination of Pap smears, and subsequent treatment of women with high-grade cervical intraepithelial neoplasia (CIN) avoids development of cancer[54].

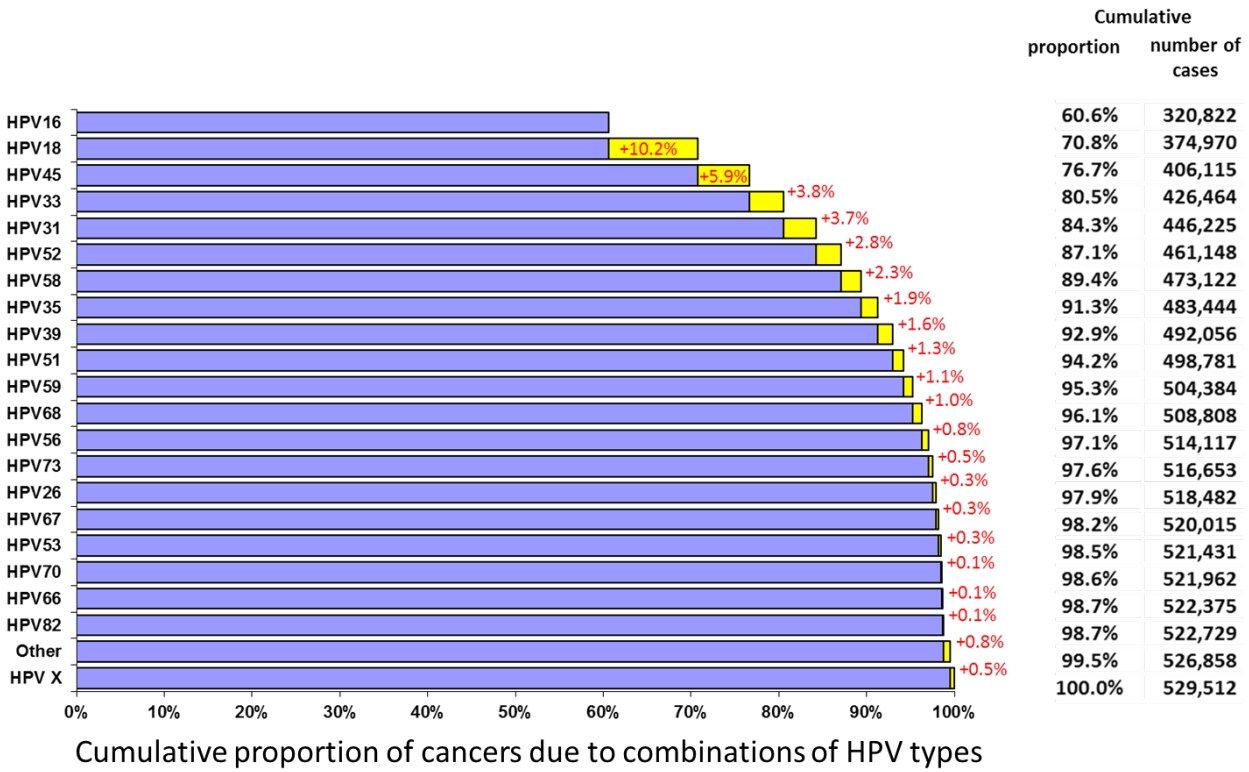


Figure 2. Cumulative proportion of cervical cancers in the World that are attributed to a ranked combination of 20 HPV types and the estimated number of cervical cancers in 2008 expected to be caused by these types. Source: Arbyn *et al*, J Pathol 2014 **[19].**

Diverse methods can be used to ablate (freezing, burning, laser) or to excise (cold knife conisation, laser conisation, large loop excision of the transformation zone) the pre-cancerous lesion[55,56]. In 1986, the effectiveness of cervical cancer screening using Pap smears was established by an expert team invited by the International Agency for Research on Cancer (IARC[55]) that pooled evidence from case-control and cohort studies, considering also differences in incidence between areas or periods with different screening coverage [57]. Further evidence has been generated from more recent studies [58-61], confirming the conclusion that well-organised cytological screening, repeated every three to five years in the age range 25-64 years may reduce the incidence of cervical cancer up to 80%[43]. However, only a few countries, such as Finland and England, in Europe and British Columbia, in Canada, have demonstrated such a level of effectiveness [58,62-65]. Cytology-based cervical cancer screening requires a comprehensive health infrastructure, skilled human resources, and quality assurance in place[66]. Given these limitations in cervical cytology, new tests are looked for, which are more reliable, less dependent on human interpretation, automatable and available at acceptable costs.

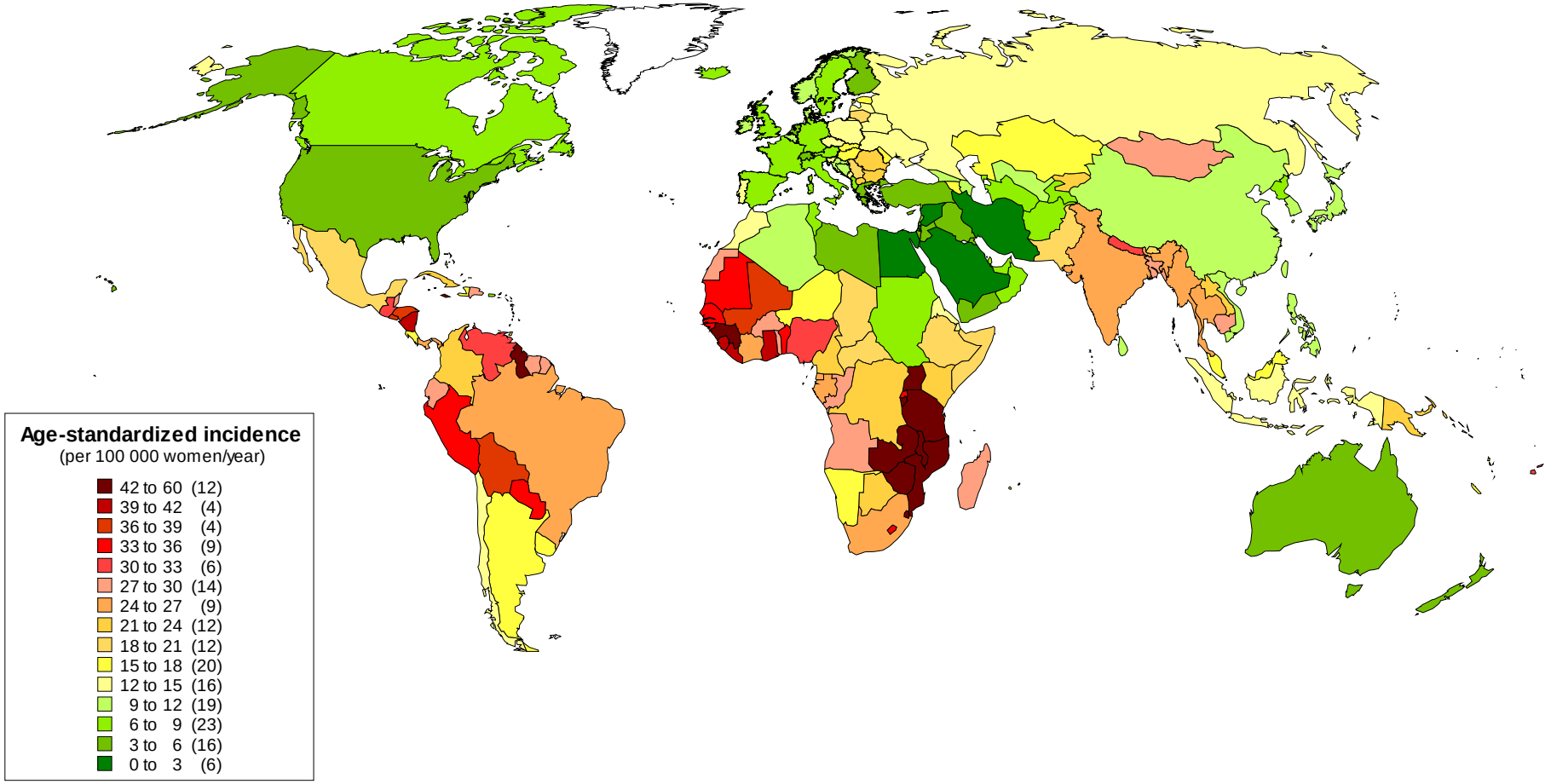


Figure 1. Geographic distribution of the world-age-standardised incidence rate (ASIR) of cervical cancer, by country, estimated for 2008 (per 100,000 women-years). The counts in brackets in the legend correspond to the number of countries in each ASIR range (Arbyn et al, *Ann Oncol* 2011)[47].

The incidence of cervical cancer varies widely among countries with world age-standardised rates ranging from <1 to >50 per 100 000. The current geographical variation in the burden of cervical cancer reflects the balance between the consequences of exposure of cohorts to HPV infection and the impact of cytological screening[47]. The lowest rates of cervical cancer are noted in areas with low prevalence of HPV (Middle-East, Maghreb, Mediterranean countries, South-Western Asia) and in areas whit successful screening programmes in the Nordic and certain West-European countries, North-America, Australia and New Zealand (Figure 3). Cervical cancer is a predominant cancer in sub-Saharan Africa, South-Eastern Asia, Melanesia, the Caribbean Islands and parts of Central and South America, which are characterised by a high prevalence of HPV and poor access to effective preventive services[47,67].

5.4. Applications of HPV testing in clinical practice

The recognition of the strong causal relationship between persistent infection of the genital tract with high-risk human papillomavirus (HPV) types and occurrence of cervical cancer[9,43] has resulted in the development of test platforms that detect HPV DNA or RNA. Detection of high-risk HPV DNA is considered to be potentially useful in three clinical applications: (1) as a triage test to select women, whose cytology is equivocal or mildly abnormal, to identify those who need referral for diagnosis and treatment; (2) as a follow-up test for women treated for cervical pre-cancer by local surgery or ablation to predict cure or failure of the treatment; and (3) as a primary screening test, solely or in combination with cervical cytology to detect cervical pre-cancer and to rule it out in the predominately healthy population[68-70]. Nowadays, particular, attention is given to the evidence regarding HPV-based primary screening as a new paradigm of cervical cancer prevention and on the identification of HPV assays, which fulfil minimal requirements, allowing use in primary cervical cancer screening. An overview of the virological test characteristics of the HPV assays can be found in Annex 15.2.

This thesis report aims summarising and updating recently conducted meta-analyses and systematic reviews which synthesise current knowledge on the performance of HPV DNA testing in each of these three clinical applications.

5.5. Background regarding HPV testing to triage women with minor cervical cytological lesions

Cervical cytology has been and still is, in most industrialised countries, the mainstay of secondary prevention of cervical cancer. Women with cytological lesions require further follow-up or treatment, or both, depending on the severity of the lesion. Women with high-grade cytological lesions should be referred immediately for further diagnostic work-up[71,72]. However, more options are considered for management of women with minor cytologic lesions[73-77].
The term minor cytological abnormalities encompass the following categories:

- Atypical squamous cells of undetermined significance (ASCUS defined according to the 1988 Bethesda System (TBS-1988)[78] and ASC-US defined according to TBS-2001)[8]. The equivalent term of the British Society of Clinical Cytology (BSCC) is: borderline dyskaryosis[79].
- Low-grade intraepithelial lesions (LSIL)[8,78] or its BSCC equivalent: mild dyskaryosis[79].

Until recently, follow-up recommendations for women with ASC-US and LSIL varied from conservative repeat cytology[80-82], to immediate referral for colposcopy and biopsy[83-85]. Although most women with an ASC-US or LSIL smear result do not have clinically significant disease, a substantial proportion of them do have histo-pathologically confirmed high-grade cervical intra-epithelial neoplasia (CIN)[86-88]. From a population of women screened in the USA, it was estimated that one third of CIN were discovered on follow-up of a previous smear with ASC-US[88]. According to previous meta-analyses, the absolute risk of underlying high-grade CIN (grade II or III or worse [CIN2/3+]), among women with ASC-US, is on average, 9-10 % for CIN2+ and 4-5 % for CIN3+[68]. For women with LSIL, these risks are about 1.5 to 2 times as high[68]. The risks are 10 to 30 times higher than for women with normal cytology.

An appropriate triage method should identify those women that have, or will develop, a cervical cancer precursor. At the same time, an accurate triage should reduce the risk of over-diagnosis and overtreatment, including adverse obstetric outcomes associated with excision of CIN lesions[89,90]. Given the evidence concerning the etiological role of high-risk human papillomavirus (hrHPV) infections in the development of cervical cancer and its precursors[1,9,91,92], HPV testing has been proposed as an alternative triage method to distinguish between women with minor cytological lesions who need referral for colposcopy, and those who can be referred back to the normal screening schedule[73,87].
In this report, an update is made of previous meta-analyses and a Cochrane review conducted at the Unit of Cancer Epidemiology of the Belgian Cancer Centre which address the accuracy of high-risk HPV testing by the Hybrid Capture assay (HC2), other hrHPV tests and repeat cytology[68,69,93,94].

5.6. Background regarding HPV testing as a test of cure after treatment for cervical pre-cancer

The presence of cervical intraepithelial neoplasia of grade two and three (CIN2-3) is associated with a risk of developing invasive carcinoma if not treated[95]. Conservative treatment of high-grade CIN, usually performed under local anaesthesia in an outpatient setting, prevents progression to cancer[96].
However, treatment is not always successful and residual or recurrent high grade disease (CIN2+) occurs in on average 8% (ranging from 4% to 18%) of treated women[69], with the majority of treatment failure occurring mainly in the first two post-operative years (Figure 4)[97-99]. Moreover, women with a history of treatment for cervical pre-cancer are still at higher risk of developing invasive cervical cancer in the future, compared to the general population during at least 10 years and maybe up to 20 years after treatment[100,101]. A recent Swedish trend study also confirmed that women with prior treatment for CIN3+ show an increased incidence of and even increased mortality from cervical or vaginal cancer in particular in older age groups (aged 60 or older) and among those treated in more recent periods[102,103]. Therefore, finding an indicator that predicts this risk with great accuracy would be particularly helpful.
Guidelines for surveillance strategies after treatment of CIN vary greatly among countries with respect to timelines, type of tests performed, and length of follow-up. In chapter 9, we update a meta-analysis on the accuracy of hrHPV testing versus cytology to detect residual/recurrent CIN2+[104,105].

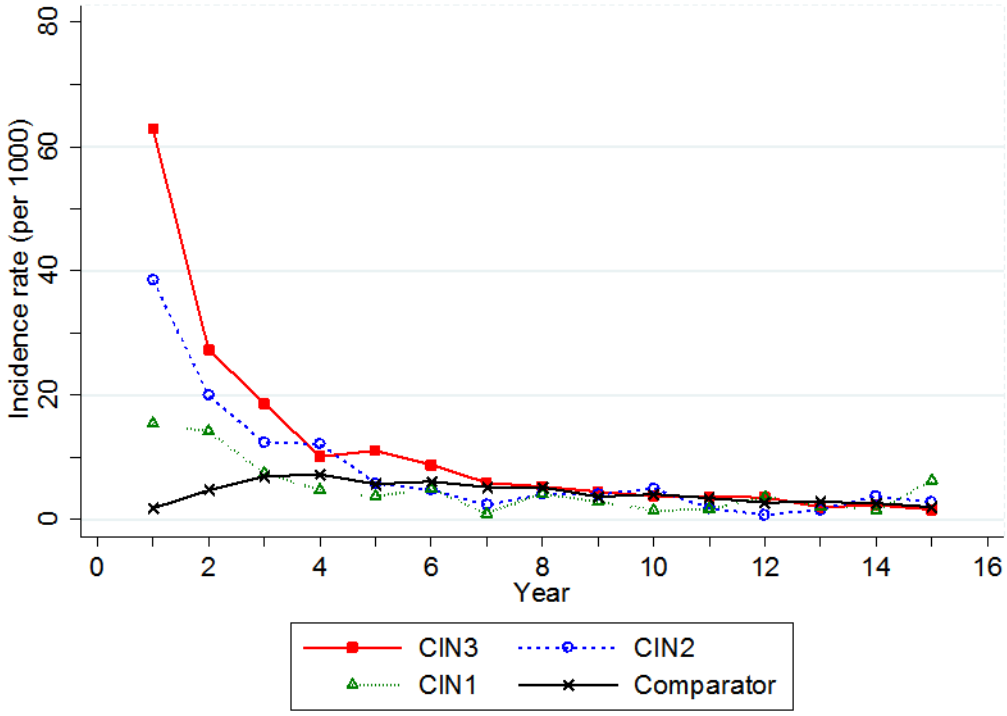


Figure 4. Incidence rates of high-grade CIN per 1000 women after treatment for CIN3 (red curve), for CIN2 (blue interrupted curve) or for CIN1 (black dotted curve). The turquoise curve represents the incidence of high-grade CIN for women with normal cytology (Comparator). Adapted from Melnikow et al., *Obstet & Gynecol* 2009[99]

5.7. Background regarding HPV testing in primary cervical cancer screening

The recognition of the strong causal relationship between persistent cervical infection with high-risk human papillomavirus (HPV) types and occurrence of cervical cancer [106] has led to the development of a series of HPV DNA or RNA tests.

Detection of high-risk (hr) HPV DNA is considered to be potentially useful as a primary screening test, solely or in combination with cervical cytology to detect cervical pre-cancer and to rule it out in the predominately healthy population. In addition new tests, based on the molecular carcinogenic pathways subsequent to expression of viral oncogenes, have been developed.

Chapter 10 updates previously conducted systematic reviews which synthesize evidence with respect to HPV-based primary screening as a new paradigm of cervical cancer prevention[69,107,108]. The chapter addresses subsequently: the accuracy of HPV tests to detect cervical pre-cancer, the efficacy of HPV screening to prevent cervical pre-cancer and cancer and the clinical validation of HPV assays. The results of systematic reviews summarised in this chapter constituted the basis for the Supplements to the 2nd Edition of the European Guidelines for Quality Assurance in Cervical Cancer Screening[109,110] and for a health technology assessment report of the Belgian Knowledge Centre[111].

Whether HPV testing is as accurate on a sample taken by a woman herself (self-sample) as on a sample taken by a clinician, to detect cervical pre-cancer, is addressed in Chapter 11. Offering a self-sampling device to women, who do not attend at regular screening, has the potential to increase screening coverage. This latter question was also covered in a meta-analysis of randomised trials conducted by the Unit of Cancer Epidemiology (Brussels)[112]. However, since the topic was related to screening participation and not to diagnostic accuracy, the latter meta-analysis was not included in this PhD manuscript.

6. Objectives

6.1. Primary objective

This thesis project aims assessing the accuracy of HPV tests in four clinical applications:

a) Triage of women with minor cytological abnormalities to identify women who need further diagnostic work-up with colposcopy and/or biopsies.
b) Follow-up of women who have been treated for a high-grade cervical intra-epithelial neoplasia in order to predict treatment success or failure.
c) Primary screening for cervical cancer.
d) HPV testing on self-samples collected by the woman her-self.

Moreover, the efficacy of screening using HPV assays to prevent cervical cancer is an additional primary objective.

6.2 Secondary objectives

By conducting systematic reviews and meta-analyses, an answer was sought for a number of specific questions related to each of the clinical applications where use of HPV tests are considered.

6.2.1. Triage of women with minor abnormal cervical cytology

- What is the clinical accuracy (sensitivity and specificity) to identify or exclude high-grade cervical pre-cancer or worse (CIN2+, CIN3+, AIS+) using hrHPV testing or biomarkers among women with a cytology result of ASC-US or LSIL?
- How compares this sensitivity and specificity with that of repeat cytology or the Hybrid Capture-2 assay?
- For newer HPV assays: what is the relative sensitivity and specificity of new HPV assays compared to the HC2 assay, which the standard comparators test?

6.2.2. Follow-up of women treated for cervical pre-cancer

- What is the sensitivity and specificity of hrHPV testing, of follow-up cytology and of combined testing with and hrHPV assays and follow-up cytology (co-testing) to predict residual or recurrent CIN2+ among women treated for high-grade cervical pre-cancer?
- What is the relative sensitivity and specificity of hrHPV testing compared to follow-up cytology and of co-testing versus hrHPV testing alone to predict residual or recurrent CIN2+ among women treated for high-grade cervical pre-cancer?

6.2.3. Primary cervical cancer screening

- Questions related to the clinical accuracy:
    - What is the absolute accuracy of HC2, the most used hrHPV DNA assay, and of other HPV tests or biomarkers to detect underlying CIN2+ or CIN3+ in women attending cervical cancer screening?
    - What is relative accuracy of HC2 compared to cervical cytology and of HC2 compared to other HPV tests or biomarkers to detect underlying CIN2+ or CIN3+ in women attending cervical cancer screening?

- Questions related to the efficacy or effectiveness of HPV-based cervical cancer screening:
    - What is the reduction of the cumulative incidence (detection rate) of CIN3+ and cervical cancer in women screened at enrolment with a hrHPV assay versus cytology and who were negative at enrolment screening?
    - What is the cumulative risk of CIN3+ or cancer after a previous negative HPV test compared to after a negative Pap test?
    - How should HPV-positive women be triaged?
    - What would be the appropriate interval to the next screening after a negative HPV test result?
    - What would be the most appropriate target age group for HPV-based screening?
    - Which HPV assays can be used as primary cervical cancer screening test?

6.2.4. HPV testing on self-samples

- What is the accuracy of HPV tests on self-samples to detect CIN2+ or CIN3+, stratified by clinical application?
- What is the relative accuracy of HPV testing on self-samples compared to on clinician-taken samples to detect CIN2+ or CIN3+?

7. Methods

7.1. Definitions and terminology

The Bethesda-2001 version for reporting of cervical cytology was used define cytological lesions of the cervix uteri epithelium[8]. Minor cytological cervical abnormalities included equivocal and low-grade cervical squamous lesions. Equivocal or atypical squamous lesions were termed as "ASC-US" (atypical squamous cells of undetermined significance). "ASCUS" (without hyphen) as defined in 1988 Bethesda System (TBS-1988)[78] and "borderline dyskaryosis"[79], the equivalent term of the British Society of Clinical Cytology (BSCC) were considered as equivalent to ASC-US. Low-grade intraepithelial lesions ("LSIL")[8,78] or its BSCC equivalent, "mild dyskaryosis"[79] was the second cytological abnormal entity considered in this PhD thesis.

The CIN nomenclature was used to describe results of the histological examination of biopsies from the cervix uteri [7].

Therapeutic or treatment failure after ablation or excision of cervical pre-cancer was defined as occurrence of a residual or recurrent CIN2+ lesion.

7.2. Meta-analysis: purpose and methods

A meta-analysis completes a systematic review where relevant information derived from all retrievable studies assessing a given clinical question is compiled[113]. A meta-analysis provides a quantified synthetic overall measure that synthesizes the data from all the selected studies. By using appropriate statistical methods, a weighted average measure can be computed that takes the inter-study variability in to account[114]. Besides producing a pooled measure, a meta-analysis must explore the inter-study heterogeneity in the study findings and try of explain the sources of this heterogeneity[115]. To answer the four clinical questions, defined in chapter 6, the PRISMA guidelines for reporting of meta-analyses and recommendations developed by the Cochrane Collaboration for systematic reviews were followed[116,117].

7.3. Literature retrieval

Several meta-analyses on the performance of various hrHPV tests, applied in triage of minor abnormal cervical cytology, as test of cure after treatment of cervical pre-cancer and in screening for cervical cancer have been conducted at the Unit of Cancer Epidemiology[69,93,94,104,108,118-123]. To update these systematic reviews, the electronic databases Pubmed Medline (http:// www.ncbi. nlm.nih.gov/sites/entrez), EMBASE (http://www.embase.com), and CENTRAL (http://www.cochranelibrary.com/about/central-landing-page.html) were consulted. Box 1, below, contains the Medline search string, used for the meta-analysis on triage of women with ASC-US or LSIL. The search strings used for the other meta-analyses can be found in the appendix (sub-chapter 15.3).

There was no language restriction.

Box 1. Medline search string used for literature retrieval regarding triage of women with ASC-US or LSIL

((cervix OR cervical) AND (cancer OR carcinoma OR neoplas* OR dysplas* OR CIN OR SIL) OR (cervix neoplasm[mesh])) AND (HPV OR (HPV AND DNA) OR (HPV AND viral) OR "human papillomavirus" OR hybrid capture OR HC2-assay OR HC2 OR HC-2 OR HCII OR HC-II) AND (triage OR management OR "follow up")

Scopus (www.scopus.com) was consulted to investigate citations of previous systematic reviews and meta-analyses on the four topics assessed in this PhD thesis. In addition, the CERVIX1 Reference Manager, a bibliographic database constructed at the Unit Cancer Epidemiology, at the Scientific Institute of Public Health (Brussels), containing more than 30,000 references on cervical cancer and HPV was interrogated using customised key words. Finally, reference lists of selected papers, were screened systematically as well.

The search to update the previous meta-analysis was limited to reports:

- on triage of women with minor cytology, published between January 01, 2012 and April 30, 2013;
- on follow-up of women treated for cervical pre-cancer and on primary screening published between January 01, 2012 and July, 1st 2013.

7.4. Eligibility of studies

7.4.1. Triage of women with minor abnormal cervical cytology

Studies were eligible to answer the questions listed in 3.2.1 if (1) cross-sectional and/or longitudinal triage data were available for women with a cytological diagnosis of ASC-US or LSIL and (2) verification with the golden standard (colposcopy and targeted biopsy, possibly completed with random biopsies and/or endo-cervical curettage) was performed on all women or women with a positive triage test in case of randomised trials.
Two types of studies were considered:

- studies with concomitant testing where all study participants are tested with the HC2 assay, other HPV tests or biomarkers, repeat cytology (if done), and the reference standard;
- randomised clinical trials (RCTs) where women with ASC-US or LSIL are randomised to HPV-based triage or repeat cytology, and where subsequently all triage-positive women are submitted to verification with the reference standard. These studies only contributed to the meta-analysis of the relative sensitivity and relative PPV. When all randomised were submitted to an outcome assessment, the trial could also contribute to the meta-analysis of the absolute accuracy.

7.4.2. Follow-up of women treated for cervical pre-cancer

Studies were eligible to answer the questions listed in 3.2.2 if (1) women were treated for histologically-confirmed CIN2 or CIN3, (2) women had a cytology and hrHPV-DNA test between three and nine months post-treatment, (3) women were followed-up for at least 18 months, and (4) colposcopy and targeted biopsy was performed on all women or women with a

positive hrHPV-DNA test or abnormal cytology. Treatment failure was defined as a follow-up histological diagnosis of CIN2+.

7.4.3. Primary cervical cancer screening

Two types of study design were considered to answer questions listed in 3.3.3: (1) cross-sectional studies where women were submitted to concomitant testing with cervical cytology (conventional or liquid), a HPV DNA assay and, optionally, other screening tests and (2) randomised clinical trials where women were assigned to cytology, HPV testing or combined testing. In the assessment of absolute sensitivity and specificity, we distinguished three situations: (1) all cases were verified with a reference standard, (2) only screen test positive cases were verified and the assumption was made that none of the women being negative for all tests had underlying CIN2+ and (3) studies, where also a random sample of women being negative for all tests were submitted to verification. For the evaluation of relative sensitivity, we considered the ratio of absolute sensitivities including intra-arm comparisons of randomised trials with combined cytology and HPV testing and the ratio of the detection rates of CIN2+ from the inter-arm comparison of RCTs. Studies were selected only when the participating women were representative of the general population.

To address the questions regarding the efficacy or effectiveness of HPV-based screening, randomised trials and screening cohort studies, respectively, were considered under the condition that they reported longitudinal outcomes, which means results at least one subsequent screening round following the enrolment screening round for women who were screen test negative at enrolment.

Studies were eligible for the question on the performance of triage algorithms to manage HPV+ women identified through screening if cross-sectional outcomes of at least one of the following tests were documented: cytology reflex testing (at cut-offs ASC-US, LSIL or HSIL), HPV genotyping for HPV16, or HPV16/18, and overexpression of p16 identified by cyto-immunochemistry. Combinations of triage tests, applied on screening samples, or applied at a subsequent visit when the first reflex triage test was negative, were also considered.

To update the list of HPV assays that fulfil requirements for use in primary cervical cancer screening, studies were considered when women were screened with multiple hrHPV assays and one of the tests was HC2 or the GP5+/6+ -EIA PCR.

7.4.4. HPV testing on self-samples

To document the performance of HPV testing on self-samples, studies were includable if the accuracy of HPV assays to detect underlying CIN2/3+ was documented for samples taken by the woman herself and also for samples taken by a clinician. Studies conducted in a primary screening setting, a high-risk setting or in a follow-up situation were all eligible.

7.5. Targeted outcomes

Occurrence of histologically confirmed cervical intra-epithelial neoplasia of grade II or worse (CIN2) and cervical intra-epithelial neoplasia of grade III or worse (CIN3+) were the main outcomes for meta-analyses assessing the diagnostic test accuracy (DTA). CIN3 comprises both

severe dysplasia and carcinoma in situ. In the PhD manuscript, adenocarcinoma in situ (AIS) was also included in "CIN3+". CIN3 or worse (CIN3, invasive squamous cervical cancer, and adenocarcinoma of the cervix (CIN3+)). CIN3 is the most relevant clinical outcome since it is considered as a precursor of cervical cancer[124]; whereas CIN2 is a mixed condition, that is less reproducible[77,125], with a lower probability of progression to cervical cancer[126,127]. However the outcome CIN2+ is reported more often than CIN3+.

Reviews assessing efficacy of HPV-based screening considered reduction of invasive cervical cancer as main outcome and reduction of CIN3+ (including adenocarcinoma in situ) as a proxy outcome.

7.6. Reference or gold standard

The following reference standards were considered as acceptable for judging on presence or absence of the target condition: colposcopy-directed punch biopsies, completed or not with random biopsies, excision biopsies by large loop excision of the transformation zone or conization, with or without endo-cervical curettage, followed by histological verification of the biopsy specimen. Cases where no biopsy was taken, because colposcopy was negative and satisfactory, were considered as being free of CIN2+. Occurrence of repeated negative cytology was in certain studies also considered as a negative disease outcome. The choice of used gold standard verification was assessed as covariate in some but not all meta-analysis.

7.7. Data extraction and management

Two authors independently checked the eligibility of references and extracted the target data for each test/outcome combination as well as covariate data. Discordances were discussed until consensus was reached.

7.8. Outcome measures

Absolute sensitivity and specificity of the evaluated index tests for CIN2+ and CIN3+ and the relative sensitivity and specificity for these outcomes of the index tests compared to comparator tests were the mean outcomes measures in the reviews assessing diagnostic test accuracy.
The HC2 assay most often was the main index test, with cytology or other HPV tests or biomarkers were the comparator tests.

The ratio of the cumulative incidence over time of CIN3+ or cancer after a negative HPV test versus after a negative comparator screening test was the outcome of efficacy or effectiveness trials and was presented as a relative risk.

7.9. Methodological quality

The methodological quality of selected studies assessing DTA was assessed using the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) check list [128,129] comprising the following items: representative spectrum, acceptable reference test, acceptable delay between tests, partial verification, differential verification, incorporation bias, blinding of reference standard, index and comparator tests, presence of relevant clinical information and reporting of uninterpretatable results.

The risk of bias in the included randomised clinical trials was evaluated using The Cochrane Collaboration's tool and the criteria specified in chapter 8 of the *Cochrane Handbook for Systematic Reviews of Interventions* [130]. This included assessment of: the randomisation method, allocation concealment; blinding of outcome assessment; reporting of incomplete outcome data, selective reporting of outcomes.

7.10. Statistical analysis

7.10.1. Absolute accuracy

The absolute sensitivity and specificity were pooled using a bivariate normal model for the logit transforms of sensitivity and specificity[131,132]. This model incorporates dichotomised test and disease outcome data (the number of true-positives, false-positives, true-negatives and false-positives) from each included study. It includes the intrinsic negative correlation between the log odds of the true- and false-positivity rates and allows for sparse data[133,134]. Summary receiver-operating-characteristics (sROC) plots were drawn to show the joint overall and study-specific sensitivity and specificity of the evaluated test. The following Stata macros were used: *metandi*[135] and *midas*[136].

7.10.2. Relative accuracy

With METADAS, a SAS macro for meta-analysis of DTA studies, the pooled relative sensitivity and specificity of the index tests versus comparator tests were assessed by including the test as a covariate in a bivariate model[137,138]. From the point estimates, ratios of sensitivity and specificity were computed.

7.10.3. Pooling of proportions or ratios

To pool proportions, a random effect model was run using *metaprop*, a statistical procedure in STATA for meta-analysis of binomial data, developed at the Unit Cancer Epidemiology (IPH, Brussels), involving Freeman-Tukey arcsine transformation to normalise variance forcing confidence intervals within the 0 - 100% margins[139]. 95% confidence intervals were computed based on the score-test statistic. Ratios of proportions or rates were pooled with the Stata command *metan*[140]. The *metaprop* and *metan* were required when the bivariate model for diagnostic accuracy did not converge because of sparse data.

7.10.4. Assessment of heterogeneity and sources of heterogeneity

Heterogeneity between studies was assessed by visual inspection of forest plots, the spread of study points in sROC plots and Cochran's *Q*-test[141]. The percentage heterogeneity between trials, which could not be ascribed to sampling variation was evaluated by the I^2 measure[142]. Where appropriate, potential sources of heterogeneity were explored by subgroup meta-analysis or meta-regression[143].

All statistical analyses were performed using STATA/SE 13 (Stata Corporation, College Station, TX, USA) and SAS 9.3(SAS Institute, Campus Drive Cary, NC, USA).

7.11. Presentation of the results

The results of the conducted meta-analyses that make up the PhD thesis are presented in the next four chapters:

- Chapter 8: Triage of women with minor abnormal cervical cytology;
- Chapter 9: Follow-up of women treated for cervical pre-cancer;
- Chapter 10: Primary cervical cancer screening;
- Chapter 11: HPV testing on self-samples.

To improve readability all tables from every chapter of the results-section are gathered at the end of each individual chapter.

Six published papers related to the conducted meta-analyses are attached in Chapter 17. The study subject addressed in each of these papers is mentioned in Table 2.

Table 2. Results of the four meta-analysis conducted in the PhD project with link to the integrated published papers.

	Chapter	8	9	10	11
	Subject	Triage abn cyto	FU after treatment	1ary screening	Self-samples
No.	Reference of papers				
1	Arbyn, J Natl Cancer Inst 2014[93]	x			
2	Arbyn, Vaccine 2012[105]	x	x	x	
3	Arbyn, Int J Cancer 2013[120]	x			
4	Arbyn, Lancet Oncol 2014				x
5	Arbyn, Cancer Epidemiol 2016[144]			x	
6	Arbyn, Ann Intern Med 2017[145]	x			

Abn cyto: minor abnormal cervical cytology; FU: follow-up.

Six Cochrane reviews related to cervical cancer prevention were not attached to the PhD manuscript because of their excessive length[94,107,146-149].

8. Triage of women with minor abnormal cervical cytology

8.1. Retrieved studies

A previously published meta-analysis on the accuracy of HC2 in triage of ASC-US[69], was updated with eleven new studies (see Figure 5). The last searching was run on April 30, 2013, which yielded in total 63 reports: 52 containing data on triage of ASC-US and 39 on triage of LSIL. The list of included studies can be found in annex 15.4.

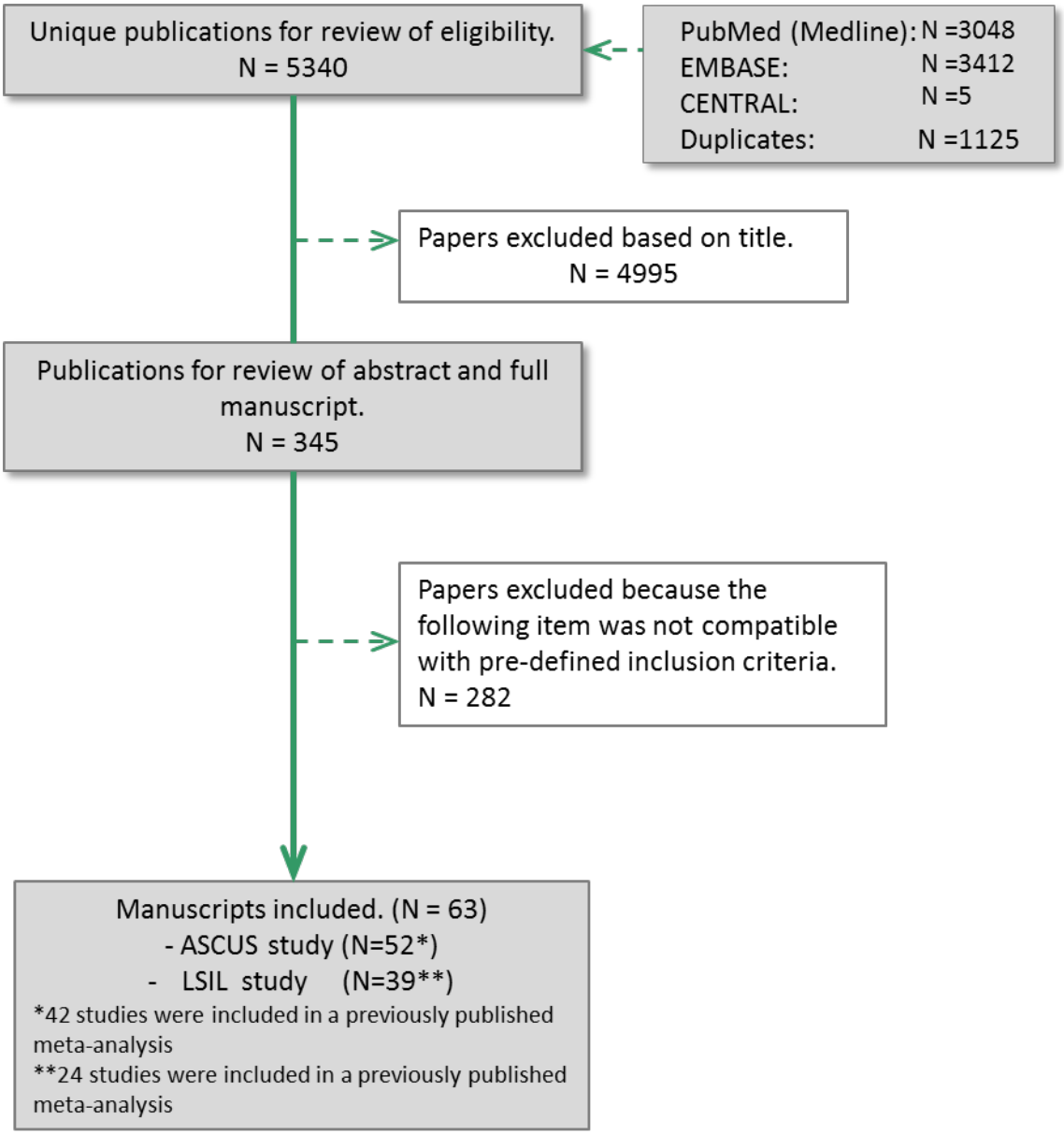


Figure 5. PRISMA flow chart for the updated retrieval of references regarding triage of women with ASC-US/LSIL (restricted to HC2). The previous literature retrieval was used for a Cochrane review[94].

The detailed quality assessment of all studies addressing the accuracy of the HC2 assay in triage of women with ASC-US and LSIL, evaluated according the QUADAS tool can be found in the Cochrane review[94].

8.2. Triage of ASC-US

8.2.1. Hybrid Capture II (HC2)

The sensitivity and specificity of HC2 used to triage women with ASC-US, to detect CIN2+, pooled from 52 studies, was 92% (95% CI= 90-93%) and 55% (95% CI= 51-60%), respectively (Figure 6 [left], Figure 7 and Table 3). HC2 detected on average 94% (95% CI= 92-96%) of CIN3+ lesions (Figure 6 [right] and Table 3).

Compared to cytology, HC2 was 27% more sensitive (relative sensitivity: 1.27 [95% CI: 1.16-1.39]) and not less specific (relative specificity: 0.99 [95% CI= 0.97-1.03]) (Figure 8 and Table 4).

Overall, the prevalence of CIN2+ and CIN3+ among women with ASC-US was 14.2% (95% CI=11.6-16.9%) and 7.8% (95% CI=5.6-10.3%), for CIN2+ and CIN3+, respectively.

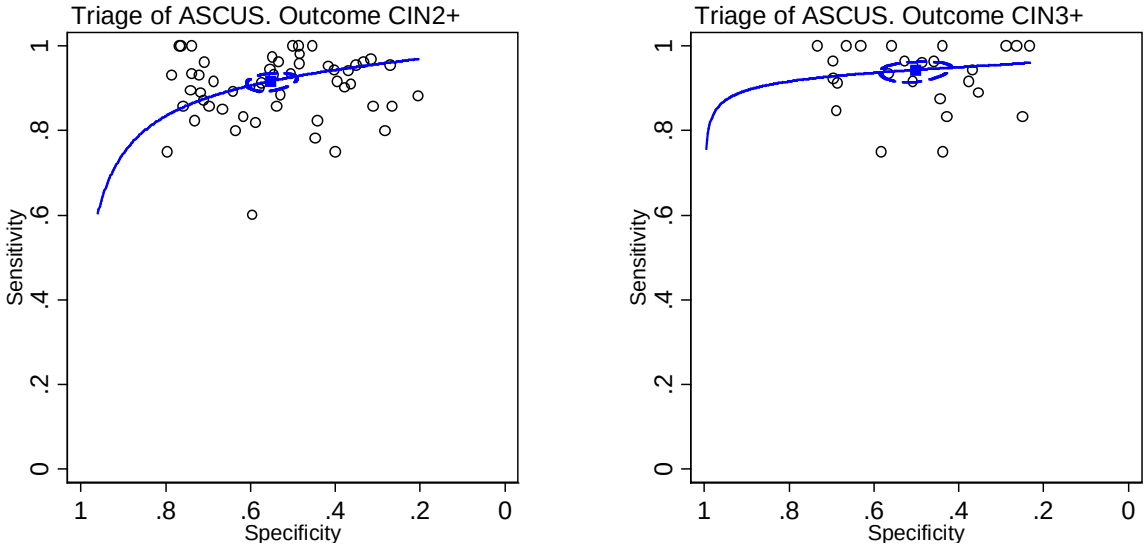


Figure 6. Meta-analysis of the accuracy of hrHPV testing with the HC2 assay to detect underlying CIN2+ (left) or CIN3+ (right) in women with ASC-US.
Black dots: study-specific sensitivity and specificity points, central square: summary point, interrupted line: 95% confidence ellipse around the summary point, full line curve: summary ROC curve.
[$] cut-off: >1.00 RLU/CO, * cut-off: ≥ASC-US

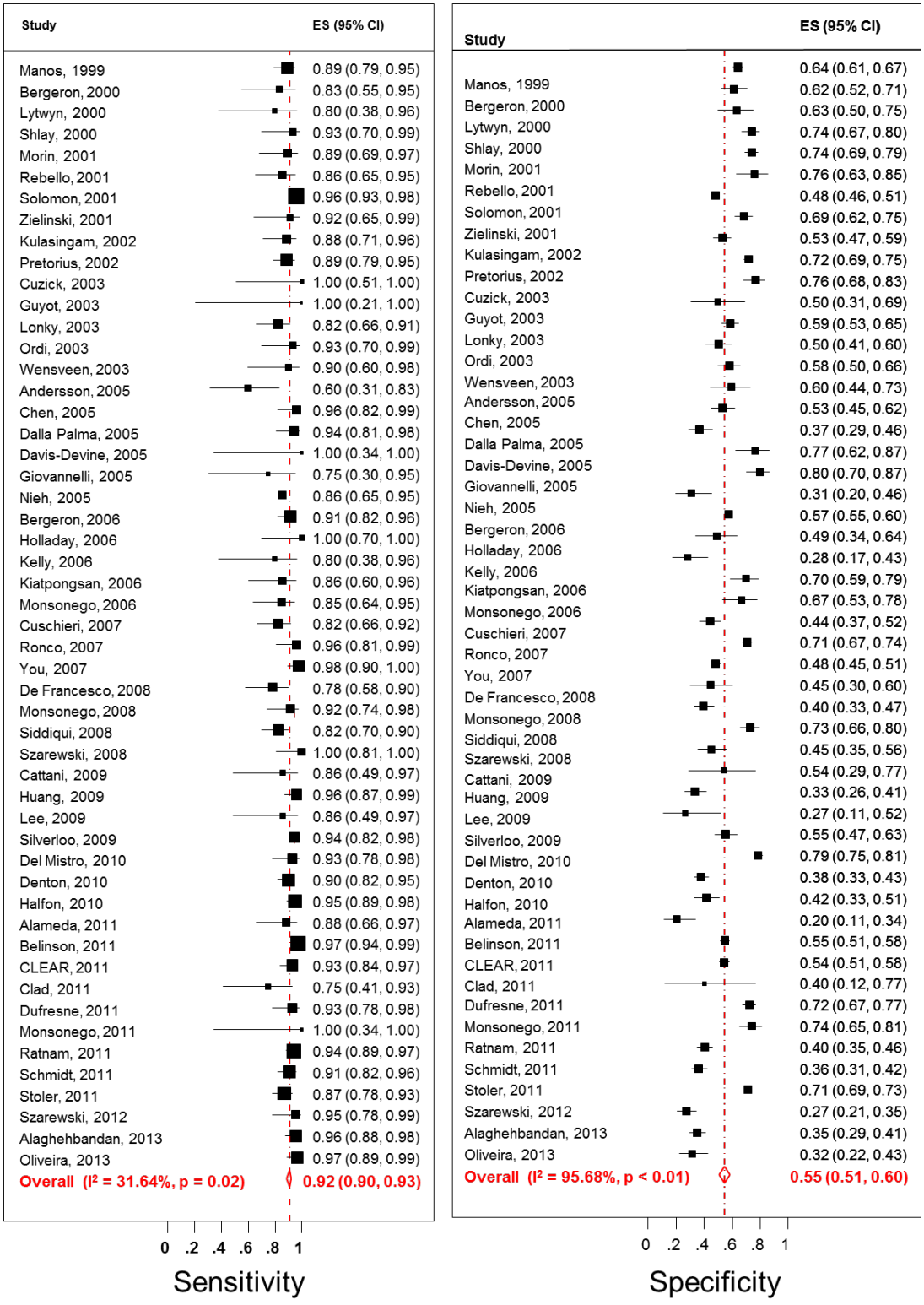


Figure 7. Meta-analysis of the sensitivity (left) and specificity (right) of hrHPV testing with the HC2 assay to detect underlying CIN2+ used to triage women with ASC-US.

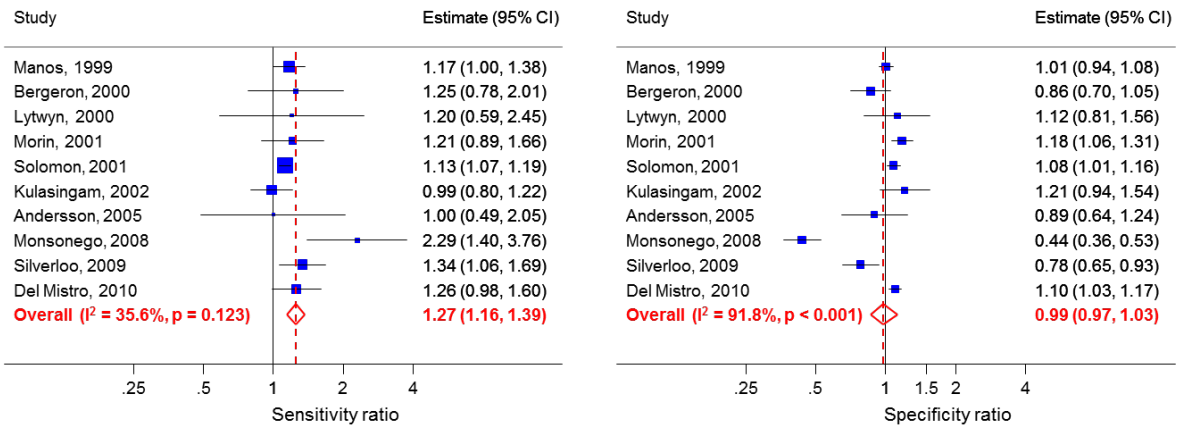


Figure 8. Meta-analysis of the relative sensitivity (left) and relative specificity (right) of repeat cytology (at cut-off ASC-US) vs HC2 (cut-off 1.00 >RLU/CO) to detect CIN2+ in women with ASC-US.

8.2.2. Other HPV assays and biomarkers

The absolute accuracy and relative accuracy to detect CIN2+ and CIN3+ of other HPV assays biomarkers was pooled and compared to HC2 (Table 6 and Table 7).

To detect CIN2+, *Linear Array, Cobas-4800, and ProExC* were found to have similar accuracy as HC2. The pooled sensitivity of the 5-type mRNA-assay, *Pretect HPV Proofer*, was significantly lower than that of HC2 (sensitivity ratio: 0.83 [95% CI=0.77-0.89]), but a significantly improved specificity was observed (specificity ratio: 1.95 [95% CI=1.68-2.25]). In the same way, *type-specific testing for HPV16 or HPV16/18* was less sensitive (ratio: 0.58 [95% CI=0.51-0.66] and 0.65 [95% CI=0.57-0.74], respectively) but more specific (ratio: 2.04 [95% CI=1.70-2.44] and 1.92 [95% CI=1.61-2.29], respectively) than testing for 13 hrHPV types with HC2.
Similar sensitivity combined with significantly improved specificity was observed for *Abbott RT-PCR* (specificity ratio: 1.18 [95% CI=1.00-1.03]), *Papillocheck* (specificity ratio: 1.44 [95% CI=1.10-1.89]), *Cervista* (specificity ratio: 1.15 [95% CI=1.06-1.24]), *APTIMA* (specificity ratio: 1.15 [95% CI=1.10-1.21]), and $p16^{INK4a}$ (specificity ratio: 1.80 [95% CI=1.38-2.34]). *Amplicor* was observed to have a similar sensitivity (ratio: 0.98 [95% CI=0.92-1.05]) but a significantly lower specificity (ratio: 0.87 [95% CI=0.79-0.95]).

8.3. Triage of LSIL

8.3.1. Hybrid Capture II

The average absolute sensitivity of HC2 to triage women with LSIL pooled from 38 studies, was 95% (95% CI=93-96%) to detect CIN2+, and the specificity was 27% (95% CI=24-31). To detect CIN3+, the sensitivity reached 98% (95% CI=95-99%) after pooling the data of 21 studies, but the specificity to exclude CIN2+ was merely 27% (95% CI=24-31%) (Table 7, Figure 9). Compared to cytology, HC2 demonstrated a 23% increase in sensitivity (ratio: 1.23

[95% CI=1.06-1.43]), but a 34% loss in specificity (ratio: 0.66 [95% CI=0.58-0.75]) to detect CIN2+ (Table 8).
Overall, the prevalence of CIN2+ and CIN3+ among women with LSIL was 21.1% (95% CI=18.0-24.5%) and 8.6% (95% CI=6.2-11.5%), respectively.

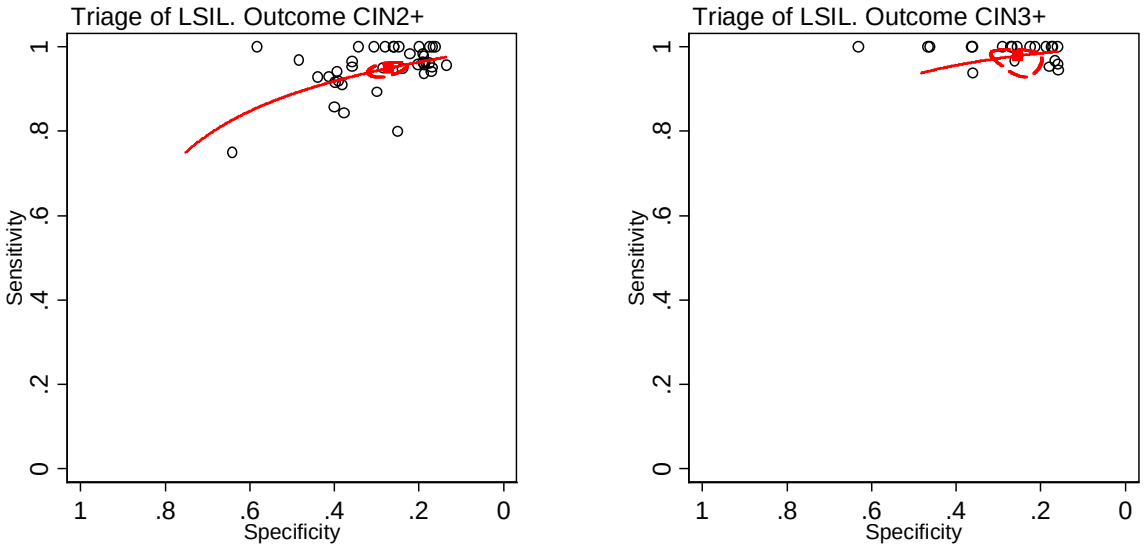


Figure 9. Meta-analysis of the accuracy of hrHPV testing with the HC2 assay to detect underlying CIN2+ (left) or CIN3+ (right) in women with LSIL.
Black dots: study-specific sensitivity and specificity points, central square: summary point, interrupted line: 95% confidence ellipse around the summary point, full line curve: summary ROC curve.

8.3.2. Other HPV assays and biomarkers

The absolute and relative accuracy to detect CIN2+ and CIN3+ of different HPV tests and biomarkers was pooled and compared to HC2 (Table 9). To detect CIN2+, the accuracy of *Linear Array*, *Papillocheck* and *Cobas-4800* was similar compared to the accuracy of HC2. The two mRNA-based assays both had a significantly improved specificity (ratio: 2.90 [95% CI=2.25-3.74] for Pretect HPV-Proofer, and 1.55 [95% Ci=1.39-1.73] for APTIMA). However, Pretect HPV-Proofer, which detects only 5 hrHPV types, had a significantly lower sensitivity (ratio: 0.74 [95% CI=0.69-0.78]) , while the 14-type APTIMA assay had a similar sensitivity (ratio: 0.97 [95% CI=0.92-1.02]) compared to HC2. A loss in sensitivity linked with an improved specificity compared to HC2 was also observed for ProExC (sensitivity ratio: 0.70 [95% CI=0.61-0.80] , specificity ratio: 3.26 [95% CI=2.47-4.29]), p16INK4a (sensitivity ratio: 0.81 [95% CI=0.69-0.95] , specificity ratio: 2.21 [95% CI=1.36-3.59]), and type-specific testing for HPV16 (sensitivity ratio: 0.57 [95% CI=0.54-0.61] , specificity ratio: 3.07 [95% CI=2.56-3.68]) and HPV1618 (sensitivity ratio: 0.62 [95% CI=0.57-0.66] , specificity ratio: 2.85 [95% CI=2.38-3.41]). Significantly improved specificity which was not linked to a loss in sensitivity for Abbott RT-PCR (sensitivity ratio: 0.97 [95% CI=0.94-1.00] , specificity ratio: 1.34 [95% CI=1.10-1.62]) and double staining with p16 and Ki-67 (sensitivity ratio: 0.94 [95% CI=0.83-1.06] , specificity ratio: 3.43 [95% CI=2.77-4.24]).

8.4. Requirements for a test usable in triage of minor cytological lesions

Preferentially, triage of screen-positive women should be performed with clinically validated tests. Experts agree on equivalency principles of minimal diagnostic test accuracy for HPV tests usable in primary screening where HC2 is accepted as a reference comparator (see guideline of Meijer et al, *Int J Cancer* 2009)[150]. Consistent evidence underpins the clinical utility of HC2 also in triage of women with equivocal squamous abnormalities. Therefore proven equivalency of an HPV assay relative to HC2 in screening, can be considered for use in triage of ASC-US. However, specific evidence regarding the diagnostic accuracy of a new triage method to predict presence of underlying or incipient squamous or glandular cervical pre-cancer in women with minor cytological abnormalities is preferable.

8.4.1. hrHPV testing in triage of ASC-US

In ASC-US triage, a new triage test should be at least as sensitive and, but preferably more specific than HC2.

→ assessing the relative accuracy of the new test versus HC2, to detect CIN2+:

- relative sensitivity: lower 95% CI bound ≥ 0.90
- relative specificity: lower 95% CI bound ≥ 0.95

Following this set of requirements the following tests are considered valid: Abbott RT-PCR, Papillocheck, Cervista, APTIMA, $p16^{INK4a}$, p16/Ki-67 (Table 11). It should be noted that the relative accuracy requirements for Papillocheck and Cervista are based on only one study. The Cobas-4800 assay demonstrates similar pooled sensitivity and specificity as HC2 but the precision of the specificity is not sufficient (lower CIB at 0.86).

8.4.2. hrHPV testing in triage of LSIL

For LSIL triage, the new test should be at least as sensitive but necessarily more specific than HC2. The requirement for higher specificity is indicated because of the very high positivity rate, and low specificity, of the latter test, in LSIL.

→ assessing the relative accuracy of the new test versus HC2, to detect CIN2+:

- relative sensitivity: lower 95% CI bound ≥ 0.90
- relative specificity: measure ≥ 1.10, and lower 95% CI bound ≥ 1.00

Following these requirements, Abbott RT-PCR and APTIMA are valid hrHPV tests in triage of women with LSIL (Table 12).

When no relative accuracy data (compared to HC2) are available, absolute accuracy verified with a good gold standard could be acceptable as well. Colposcopy and multiple biopsies of colposcopically suspicious areas and high-quality standards for colposcopy and histology, including review of biopsies are acceptable criteria.

Absolute accuracy data of a new HPV assay or biomarker can then be compared with the pooled absolute accuracy data for HC2 applied in triage of a given lesion in other studies.

8.5. Tables: Triage of women with minor abnormal cervical cytology

Table 3: Pooled absolute accuracy of HC2 and cytology in the triage of women with ASC-US in order to detect underlying CIN2+ or CIN3+.

TEST	CIN2+			CIN3+		
	No studies	**Sensitivity** (95% CI)	**Specificity** (95% CI)	**No studies**	**Sensitivity** (95% CI)	**Specificity** (95% CI)
HC2[$]	52	91.7 % (89-93.2%)	55.2% (50.5-59.8%)	25	94.3% (92.2-95.9%)	50.3 (43.9-56.6%)
Cytology*	15	74.3 (67.7-79.9)	64.3% (57.4-70.7%)	4	77.5 (63.8.1-87.9%)	68.9% (47.7-84.3%)

Table 4. Pooled relative accuracy of HC2 versus cytology in the triage of women with ASC-US to detect underlying CIN2+ or CIN3+[4] (see Arbyn et al, 2013[94]).

Test comparison	CIN2+			CIN3+		
	No. studies	**Relative sensitivity** (95% CI)	**Relative specificity** (95% CI)	**No. studies**	**Relative sensitivity** (95% CI)	**Relative specificity** (95% CI)
HC2[$] vs. cytology*	10	1.27 (1.16-1.39)	0.99 (0.97-1.03)	4	1.14 (1.06-1.22)	0.99 (0.89-1.09)

[$] cut-off: 1.00 >RLU/CO, * cut-off: ≥ASC-US

[4] The estimates of the relative accuracy of HC2 compared to repeat cytology were derived from a recent Cochrane review[94] since no new studies were identified.

Table 5. Pooled absolute and relative (compared to HC2) sensitivity and specificity of other HPV assays and biomarkers to detect CIN2+ in the triage of ASC-US.

Test	Absolute accuracy			Relative accuracy		
	No studies	Pooled sensitivity	Pooled specificity	No studies	Pooled sensitivity	Pooled specificity
Amplicor	6	91.1 (86.4-95.0)	47.2 (37.3-57.3)	4	0.98 (0.92-1.05)	0.87 (0.79-0.95)
Abbott PCR	4~	94.6 (89.4-98.3)	42.2 (34.8-49.9)	4	0.98 (0.93-1.03)^	1.18 (1.00-1.03)
Linear Array	11	94.8 (90.5-98.0)	42.4 (34.2-50.7)	5	1.02 (0.99-1.06)	0.90 (0.79-1.03)
Papillocheck[¥]	1	96.1 (92.4-99.9)	60.2 (51.0-69.4)	1	1.01 (0.95-1.07)	1.44 (1.10-1.89)
Cervista[¥]	3	95.9 (93.1-98.7)	49.6 (34.6-64.6)	1	0.98 (0.95-1.02)	1.15 (1.06-1.24)
Cobas-4800	3	91.2 (85.4-95.8)	57.8 (40.4-74.3)	2	1.02 (0.94-1.11)	1.06 (0.85-1.32)
Pretect	11	75.5 (66.4-83.6)	76.6 (66.0-85.9)	6	0.83 (0.77-0.89)	1.95 (1.68-2.25)
APTIMA	10	95.4 (91.6-98.4)	58.0 (49.3-66.4)	9	1.01 (0.97-1.04)^	1.15 (1.10-1.21)
HPV16	23	53.1 (48.3-57.9)	86.2 (83.1-88.9)	12[£]	0.58 (0.51-0.66)	2.04 (1.70-2.44)
HPV16/18	25	57.7 (52.2-63.1)	82.4 (79.0-85.6)	12[£]	0.65 (0.57-0.74)	1.92 (1.61-2.29)
ProExC	5	89.8 (68.2-100)	80.5 (73.9-86.3)	3	0.97 (0.66-1.42)	1.77 (0.92-3.42)
$p16^{INK4a}$	19[$]	84.7 (77.7-90.9)	67.6 (59.2-75.4)	8	0.98 (0.91-1.06)^	1.80 (1.38-2.34)
p16/Ki-67	4	80.4 (59.7-95.5)	65.7 (51.5-78.7)	1	1.01 (0.92-1.12)	2.22 (1.89-2.62)

~The study of Wong et al., 2011[151] was excluded due to its outlying low specificity for the Abbott PCR (20.8%). Inclusion of Wong et al., 2011 resulted in a pooled sensitivity and specificity of 95.4% (95% CI=90.6-98.8) and 38.8% (95% CI=29.1-48.0), respectively.

[¥]derived from a recently published meta-analysis since no new studies identified [69].

[£]12test comparisons from 7 studies.

[$]Denton et al. (2010)[152] reported the results of 3 independent $p16^{INK4a}$ tests (2 performed by 2 different pathologists and 1 performed by a cytotechnologist).

^continuity correction performed for studies with relative accuracy of 100% over 100%

Table 6. Pooled absolute and relative (compared to HC2) sensitivity and specificity of other HPV assays and biomarkers to detect CIN3+ in the triage of ASC-US.

Test	Absolute accuracy			Relative accuracy		
	No studies	Pooled sensitivity	Pooled specificity	No studies	Pooled sensitivity	Pooled specificity
Amplicor	3	93.1 (85.6-98.3)	37.8 (26.8-49.6)	3	1.00 (0.92-1.09)^	0.86 (0.69-1.06)
Abbott PCR	4~	96.3 (93.4-98.6)	29.4 (22.0-37.4)	3	0.99 (0.92-1.06)^	1.17 (0.97-1.41)
Linear Array	7	99.2 (97.4-100)	39.3 (30.8-48.1)	4	1.03 (0.98-1.09)^	0.91 (0.79-1.06)
Papillocheck	-	-	-	-	-	-
Cervista	2	98.0 (94.6-99.9)	50.7 (35.7-65.6)	-	-	-
Cobas-4800	2	96.1 (88.8-100)	51.6 (18.2-84.2)	2	1.01 (0.93-1.11)^	1.06 (0.88-1.28)
Pretect	8	86.0 (75.1-94.6)	75.2 (61.2-86.9)	4	0.88 (0.81-0.96)	2.08 (1.77-2.46)
APTIMA	11	98.9 (96.1-100)	53.7 (44.9-62.4)	9	1.00 (0.96-1.04)^	1.15 (1.10-1.21)
HPV16	18	65.2 (58.1-72.1)	82.0 (77.6-86.1)	9[£]	0.72 (0.62-0.84)	2.13 (1.58-2.88)
HPV16/18	19	71.5 (63.5-79.0)	78.5 (74.0-82.6)	9[£]	0.76 (0.66-0.88)	2.06 (1.55-2.74)
ProExC	3	100 (95.5-100)	78.0 (58.6-92.7)	1	1.00 (0.87-1.15)^	0.95 (0.81-1.11)
$p16^{INK4a}$	10[$]	85.3 (76.6-92.6)	62.2 (59.2-65.2)	4	1.00 (0.91-1.10)^	1.71 (1.36-2.15)
P16/Ki-67	1	75.0 (30.1-95.4)	61.1 (51.9-69.6)	-	-	-

~The study of Wong et al., 2011[151] was excluded due to its low specificity. Inclusion of Wong et al., 2011 resulted in a sensitivity and specificity of 98.9% (95% CI=92.3-100) and 37.4% (95% CI=29.4-45.7), respectively.

[£]9 test comparisons from 5 studies.

[$]Denton et al. (2010)[152] reported the results of 3 independent $p16^{INK4a}$ tests (2 performed by 2 different pathologists and 1 performed by a cytotechnologist).

^continuity correction performed for studies with relative accuracy of 100% over 100%

Table 7. Pooled absolute accuracy of HC2 and cytology to triage women with LSIL.

TEST	CIN2+			CIN3+		
	No studies	**Sensitivity** (95% CI)	**Specificity** (95% CI)	**No studies**	**Sensitivity** (95% CI)	**Specificity** (95% CI)
HC2[$]	39	94.9 (93.4-96.1)	27.2 (23.7-30.9)	21	97.9 (94.8-99.1)	25.4 (21.2-30.2)
Cytology*	9	83.1 (69.7-91.3)	47.3 (34.9-60.0)	4	80.6 (64.9-90.3)	48.9 (25.5-72.8)

[$] cut-off: >1.00 RLU/CO, * cut-off: ≥ASC-US

Table 8. Pooled relative accuracy of HC2 versus cytology to triage women with LSIL in order to detect underlying CIN2+ or CIN3+[5].

Test comparison	CIN2+			CIN3+		
	No studies	**Relative sensitivity** (95% CI)	**Relative specificity** (95% CI)	**No studies**	**Relative sensitivity** (95% CI)	**Relative specificity** (95% CI)
HC2[$] vs. cytology*	6	1.23 (1.06-1.43)	0.66 (0.58-0.75)	4	1.15 (0.89-1.38)	0.56 (0.37-0.84)

[$] cut-off: >1.00 RLU/CO, * cut-off: ≥ASC-US

[5] The estimates of the relative accuracy of HC2 compared to repeat cytology were derived from a recent Cochrane review[94] since no new studies were identified.

Table 9. Pooled absolute and relative (compared to HC2) sensitivity and specificity of other HPV assays and biomarkers to detect CIN2+ in the triage of LSIL.

Test	Absolute accuracy			Relative accuracy		
	No studies	Pooled sensitivity	Pooled specificity	No studies	Pooled relative sensitivity	Pooled relative specificity
Amplicor	4	94.1 (82.8-99.9)	29.5 (20.3-39.5)	2	0.98 (0.90-1.07)	0.80 (0.65-0.98)
Abbott PCR	4	96.3 (93.4-98.6)	29.4 (22.0-37.4)	4	0.97 (0.94-1.00)	1.34 (1.10-1.62)
Linear Array	8	99.3 (97.3-100)	26.4 (21.3-31.7)	5	1.00 (0.97-1.02)^	1.03 (0.84-1.25)
Papillocheck[¥]	2	94.6 (79.4-100)	32.1 (19.9-44.4)	2	1.05 (0.93-1.19)	0.90 (0.54-1.51)
Cervista[¥]	1	96.8 (94.0-99.6)	46.7 (40.6-52.8)	-	-	-
Cobas-4800	2	93.9 (89.4-97.3)	26.6 (17.1-37.3)	1	1.02 (0.96-1.07)	1.26 (0.98-1.62)
Pretect	10	70.1 (63.4-76.4)	73.3 (68.3-78.0)	6	0.74 (0.69-0.78)	2.90 (2.25-3.74)
APTIMA	8	96.3 (92.2-99.2)	39.6 (31.4-48.2)	7	0.97 (0.92-1.02)	1.55 (1.39-1.73)
HPV16	22	59.8 (46.2-53.5)	82.6 (79.6-85.5)	13[*]	0.57 (0.54-0.61)	3.07 (2.56-3.68)
HPV16/18	21	55.1 (51.5-58.7)	78.4 (74.6-82.0)	13[*]	0.62 (0.57-0.66)	2.85 (2.38-3.41)
ProExC	4	68.6 (60.7-76.1)	70.4 (61.2-78.9)	1	0.70 (0.61-0.80)	3.26 (2.47-4.29)
$p16^{INK4a}$	15[$]	81.3 (72.5-88.9)	59.7 (48.3-70.5)	7	0.81 (0.69-0.95)	2.21 (1.36-3.59)
$p16^{INK4a}$/Ki-67	5	90.8 (84.4-95.8)	56.5 (44.6-68.0)	2	0.94 (0.83-1.06)	3.43 (2.77-4.24)

[¥]No new studies were identified than those included in a recent meta-analysis[69]. Therefore the results are derived from the Cochrane review where a binormal model for pooling of accuracy data was used.

[*]13 test comparisons from 6 studies.

[$]Denton et al. (2010)[152] reported the results of 3 independent $p16^{INK4a}$ tests (2 performed by 2 different pathologists and 1 performed by a cytotechnologist).

^ continuity correction performed for studies with relative accuracy of 100% over 100%

Table 10. Pooled absolute and relative (compared to HC2) sensitivity and specificity of other HPV assays and biomarkers to detect CIN3+ in the triage of LSIL.

Test	Absolute accuracy			Relative accuracy		
	No studies	Pooled sensitivity	Pooled specificity	No studies	Pooled sensitivity	Pooled specificity
Amplicor	2	97.3 (76.9-100)	22.3 (12.5-33.9)	2	1.00 (0.96-1.04)^	0.80 (0.65-0.99)
Abbott PCR	3	99.8 (96.7-100)	23.5 (16.8-31.0)	3	1.00 (0.97-1.03)^	1.33 (1.07-1.65)
Linear Array	5	100 (100-100)	22.0 (15.1-29.9)	4	1.01 (0.98-1.04)^	1.05 (0.84-1.33)
Papillocheck	-	-	-	-	-	-
Cervista	1	97.2 (92.1-99.0)	39.9 (34.5-45.4)	-	-	-
Cobas-4800	2	97.4 (84.0-100)	24.7 (14.9-36.1)	1	1.00 (0.96-1.04)^	1.21 (0.95-1.56)
Pretect	7	78.8 (68.6-87.6)	69.3 (64.2-74.3)	4	0.81 (0.75-0.88)	3.27 (2.59-4.13)
APTIMA	7	100 (99.2-100)	36.0 (28.5-43.8)	7	0.99 (0.97-1.02)^	1.47 (1.26-1.70)
HPV16	17	63.3 (58.0-68.5)	78.8 (75.3-82.0)	8*	0.67 (0.62-0.73)	3.61 (3.00-4.35)
HPV16/18	18	69.7 (63.0-76.1)	76.0 (71.9-79.9)	8*	0.70 (0.65-0.75)	3.33 (2.78-4.00)
ProExC	2	91.4 (63.0-100)	68.8 (63.0-74.2)	-	-	-
$p16^{INK4a}$	8$	91.2 (83.4-97.1)	47.5 (38.4-56.7)	3	0.90 (0.79-1.02)	2.72 (2.17-3.42)
$p16^{INK4a}$/Ki-67	2	97.6 (90.7-100)	41.1 (27.1-55.9)	-	-	-

*8 test comparisons from 46 studies.
^continuity correction performed for studies with relative accuracy of 100% over 100%
$Denton et al. (2010)[152] reported the results of 3 independent $p16^{INK4a}$ tests (2 performed by 2 different pathologists and 1 performed by a cytotechnologist).

Table 11: Pooled relative accuracy of hrHPV tests and biomarkers versus HC2 (cut-off: ≥1.00 RLU/CO) considering CIN2+. These tests fulfil the requirements of a triage test for women with ASC-US.

Test	Relative sensitivity	Relative specificity
Abbott PCR	0.98 (0.93-1.03)^	1.18 (1.00-1.03)
Papillocheck	1.01 (0.95-1.07)	1.44 (1.10-1.89)
Cervista	0.98 (0.95-1.02)	1.15 (1.06-1.24)
APTIMA	1.01 (0.97-1.04)^	1.15 (1.10-1.21)
$p16^{INK4a}$	0.98 (0.91-1.06)^	1.80 (1.38-2.34)
P16/Ki-67	1.01 (0.92-1.12)	2.22 (1.89-2.62)

Table 12. Pooled relative accuracy of hrHPV tests versus HC2 (cut-off: ≥1.00 RLU/CO) considering CIN2+. These tests fulfil the requirements of a triage test for women with LSIL.

Test	Relative sensitivity	Relative specificity
Abbott PCR	0.97 (0.94-1.00)	1.34 (1.10-1.62)
APTIMA	0.97 (0.92-1.02)	1.55 (1.39-1.73)

9. Follow-up of women treated for cervical pre-cancer

9.1. Retrieved studies

The last search was run on July 1, 2013. Three studies[153-155] were found eligible and were added to the list of studies that had been included in the previously published meta-analysis[69], resulting in a total of 17 studies. A PRISMA flow chart for the retrieval of studies is shown in Figure 10. Overall, included studies were heterogeneous with regard to their study characteristics.
All included studies were prospective, except two that had a case-control design[156,157]. HrHPV-DNA testing was performed by the HC2 in ten studies[97,154,155,158-163], by a PCR method in eight studies[153,156,157,164-168]. Overall, follow-up data of 3,041 women treated for CIN2+ or CIN3+ were included in the analysis.
The quality of included studies was evaluated using the QUADAS-2 tool and is summarized in Table 13. Overall, studies scored well for the majority of the QUADAS-items, except that the blinding of disease outcome versus the evaluated tests (hrHPV-DNA testing and follow-up cytology) not assured and was in certain studies or not sufficiently documented.

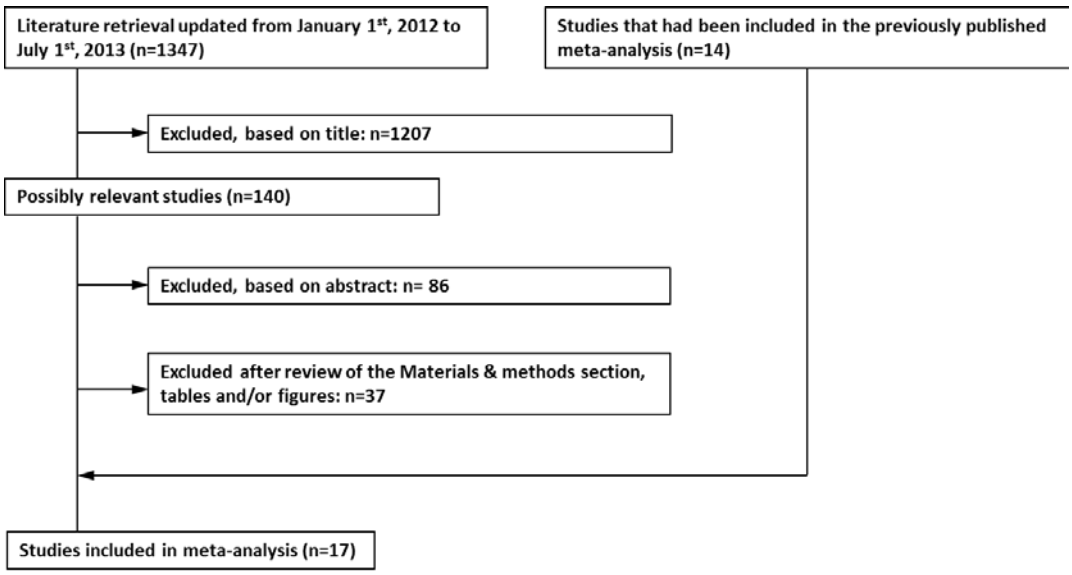


Figure 10. PRISMA flow chart for the retrieval of studies**.**

9.2. Treatment failure

Treatment failure was expressed in terms of residual or recurrent CIN2+, confirmed by histology during follow-up. Across studies with a prospective design, residual/recurrent disease was detected in 224 out of a total of 2,930 women. The proportion of CIN2+, detected during follow-up, ranged from 4%[163] to 16%[164], with a pooled estimate of 7.5% (95% CI: 6.1-8.9%), as shown in Figure 11.

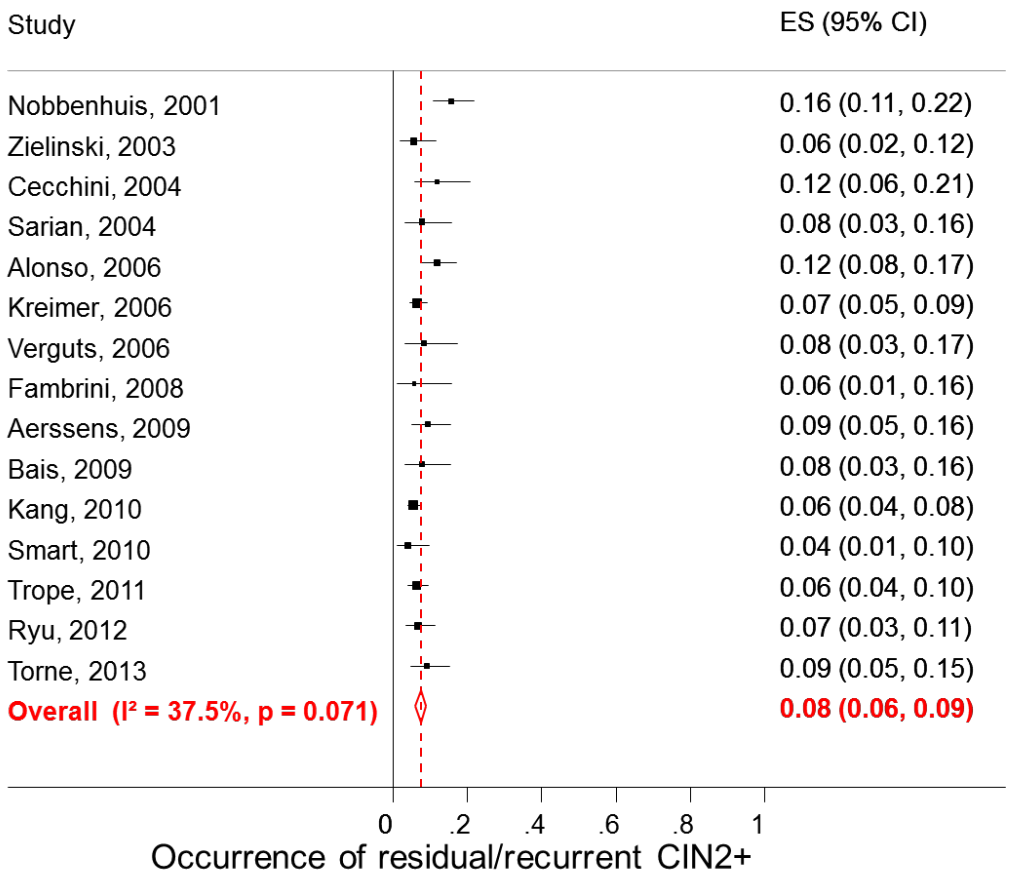


Figure 11. Meta-analysis of the occurrence of residual or recurrent CIN2+ after treatment for CIN2+ or CIN3+.

9.3. Absolute accuracy of hrHPV-DNA testing to predict occurrence of CIN2+ after treatment for pre-cancer.

Overall, the pooled sensitivity and specificity of hrHPV-DNA testing to predict occurrence of CIN2+ was 94.3% (95% CI: 88.4-97.3%) and 80.0% (95% CI: 74.2-84.8%), respectively (Figure 12 and Figure 13).

Among studies that used HC2[97,154,155,158-163], sensitivity ranged from 83% to 100.0%, and specificity ranged from 64% to 93.% (Figure 13). The pooled sensitivity and specificity of HC2 to detect residual/recurrent CIN2+ was 95.6% (95% CI: 89-7-98.2%) and 81.9% (95% CI: 75.3-87.1%), respectively. Across all eight studies that used a PCR-based test[153,156,157,164-168], sensitivity and specificity ranged from 54% to 100% and 57% to 100%, respectively (Figure 13). The pooled sensitivity and specificity for PCR-based assays were 93.1% (95% CI=86.3-99.9%) and 76.9% (95% CI=67.8-86.1%), respectively.

In the study of Aerssens et al.[167], the lowest sensitivity for hrHPV testing was observed although the analytically highly sensitive SPF10-PCR assay is used. Performing a sensitivity analysis by excluding this study, resulted in a pooled sensitivity and specificity of PCR-based testing of 95.8% (95% CI=90.9-1.00%) and 76.5% (95% CI=65.8-87.1%), respectively.

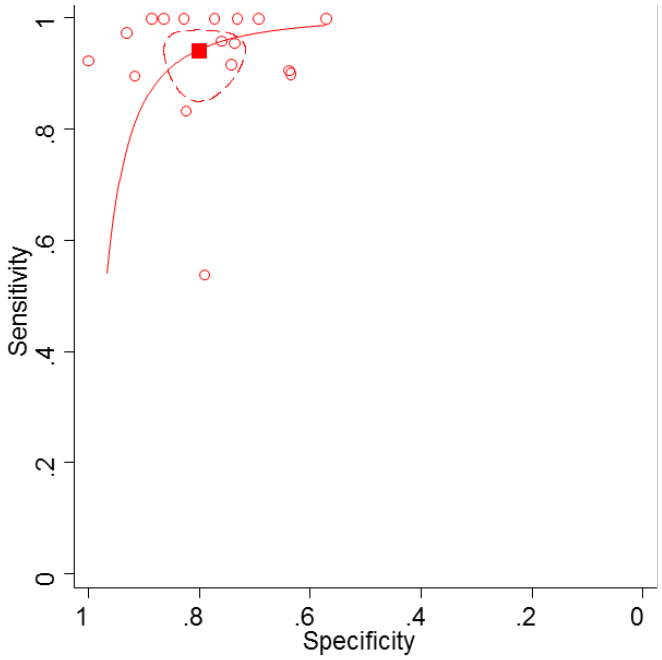


Figure 12. Specificity and sensitivity of hrHPV-DNA testing to predict occurrence of CIN2 or CIN3 after treatment of cervical pre-cancer. Full line= summary ROC curve, Square= pooled summary measure, interrupted line is the 95% confidence ellipse around the summary measure.

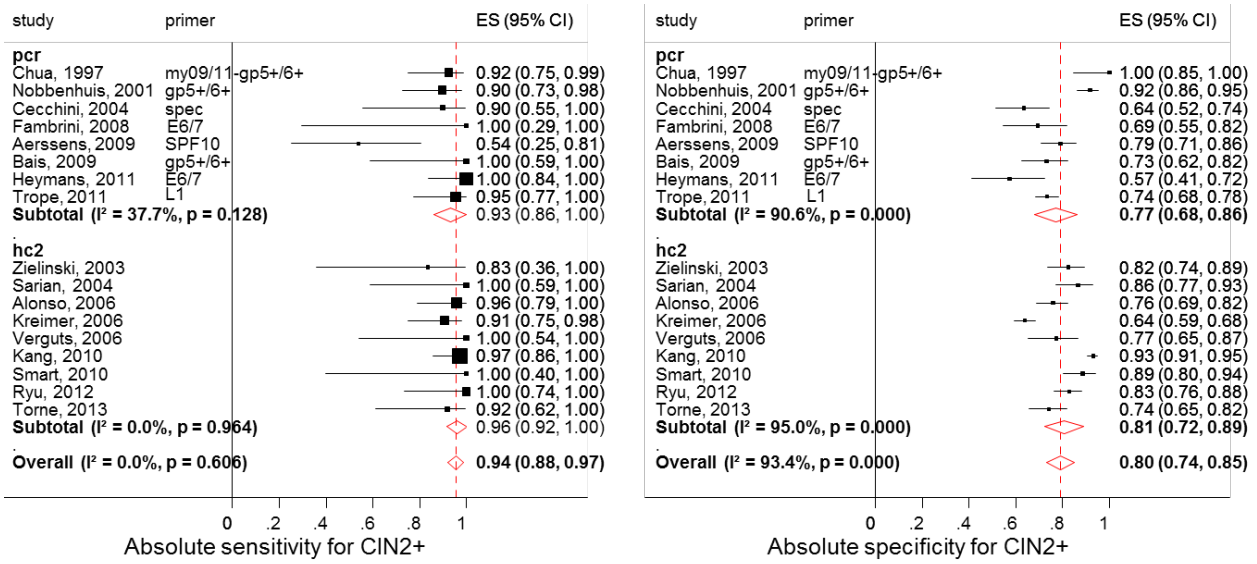


Figure 13. Study-specific and pooled absolute sensitivity (left) and specificity (right) of hrHPV-DNA testing by HC2 and PCR to predict occurrence of CIN2+ after treatment for CIN2+ or CIN3+.

9.4. Absolute accuracy of cytology follow-up to predict residual/recurrent CIN2+

The pooled absolute sensitivity and specificity to predict occurrence of CIN2+ after treatment for CIN2+ was 72.0% (95% CI: 65.6-77.5%; range: 50.0-100.0%) and 84.6% (95% CI: 80.7-87.9%; range: 64.3-91.8%), respectively) (see Figure 14).

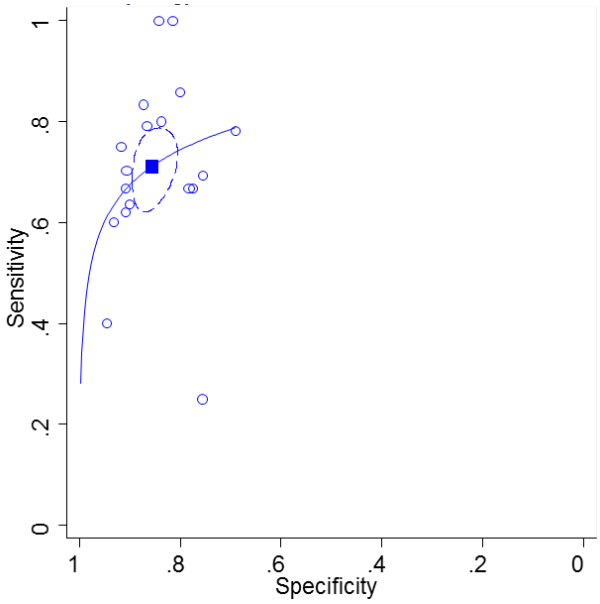


Figure 14. Specificity and sensitivity of cytology (cut-off: ASC-US) to predict occurrence of CIN2 or CIN3 after treatment of cervical pre-cancer. Full line= summary ROC curve, Square= pooled summary measure, interrupted line is the 95% confidence ellipse around the summary measure.

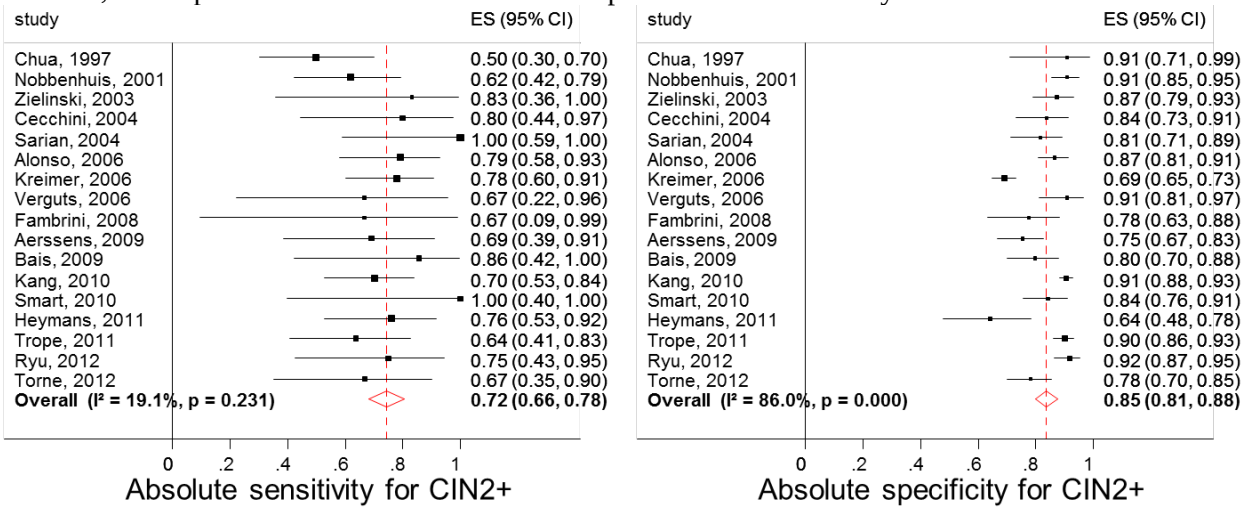


Figure 15. Study-specific and pooled absolute sensitivity (left) and specificity (right) of cytology (cut-off ASC-US) to predict occurrence of CIN2+ after treatment for CIN2+ or CIN3+.

9.5. Absolute accuracy of combined hrHPV-DNA and cytology testing to predict occurrence of CIN2+ after treatment for pre-cancer

Ten studies contained accuracy data on combined testing with an HPV-DNA assay and cytology to detect residual/recurrent disease. Pooled values for sensitivity and specificity were 95.3% (95% CI: 88.1-98.2%) and 69.6% (95% CI: 61.7-76.5%), respectively (see Figure 16 and Figure 17).

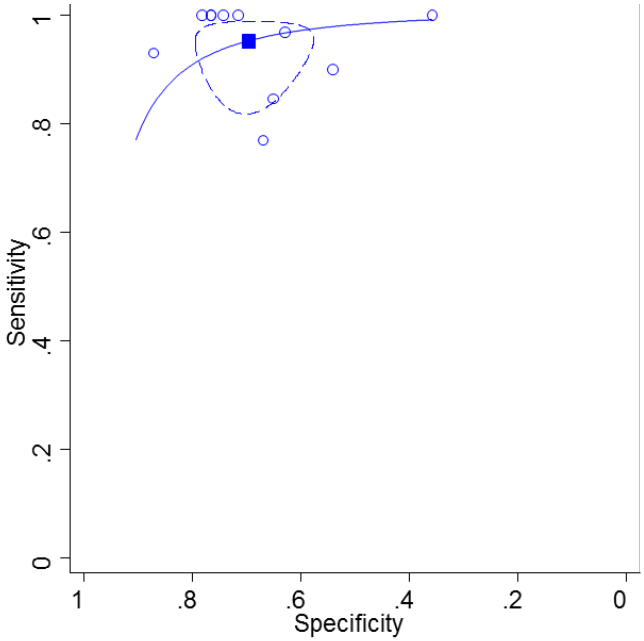


Figure 16. Specificity and sensitivity of combined testing with hrHPV-DNA testing and cytology (cut-off: ASC-US) to predict occurrence of CIN2+ after treatment of cervical pre-cancer. Full line= summary ROC curve, Square= pooled summary measure, interrupted line is the 95% confidence ellipse around the summary measure.

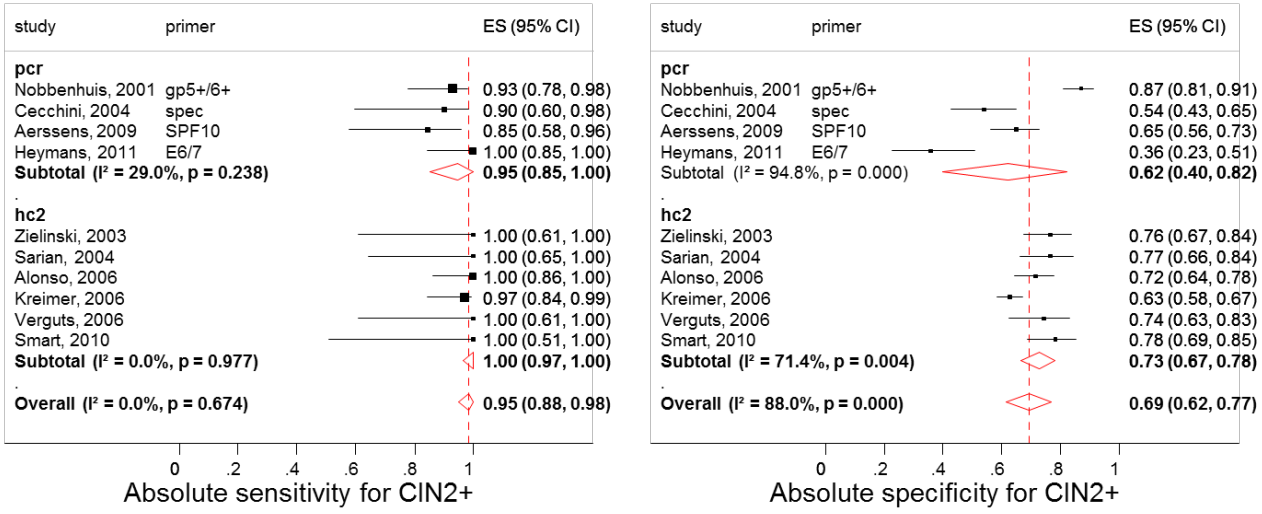


Figure 17. Study-specific and pooled absolute sensitivity (left) and specificity (right) of combined testing with an hrHPV-DNA assay (PCR, on top, or HC2, below) and cytology (cut-off: ASC-US), to predict occurrence of CIN2+ after treatment for CIN2+ or CIN3+.

9.6. Relative accuracy of hrHPV-DNA testing versus cytology to predict occurrence of CIN2+ after treatment for pre-cancer.

Compared to cytology, hrHPV-DNA testing at 3-9 months post-treatment was significantly more sensitive, but significantly less specific to predict residual/recurrent CIN2+. The pooled values for relative sensitivity and specificity were 1.29 (95% CI: 1.18-1.40) and 0.94 (95% CI: 0.90-0.99), respectively (Figure 18 and Figure 19).

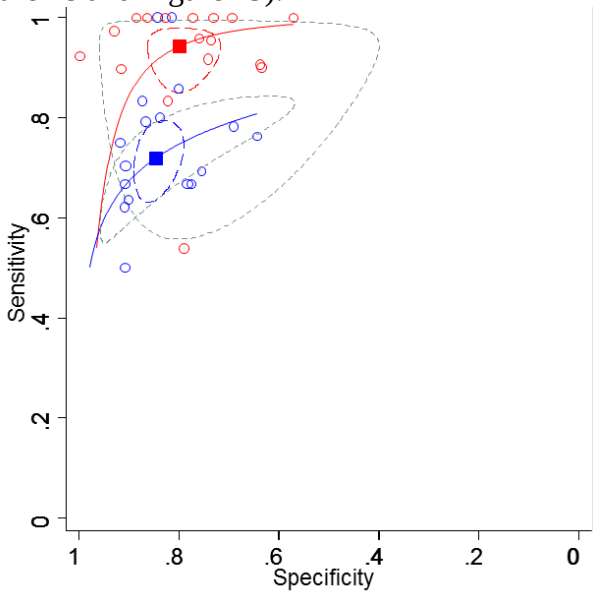


Figure 18. Specificity and sensitivity of hrHPV-DNA testing (red) and cytology at cut-off ASC-US (blue) to predict occurrence of CIN2+ after treatment of cervical pre-cancer. Full line= summary ROC curve, Square= pooled summary measure, interrupted line is the 95% confidence ellipse around the summary measure.

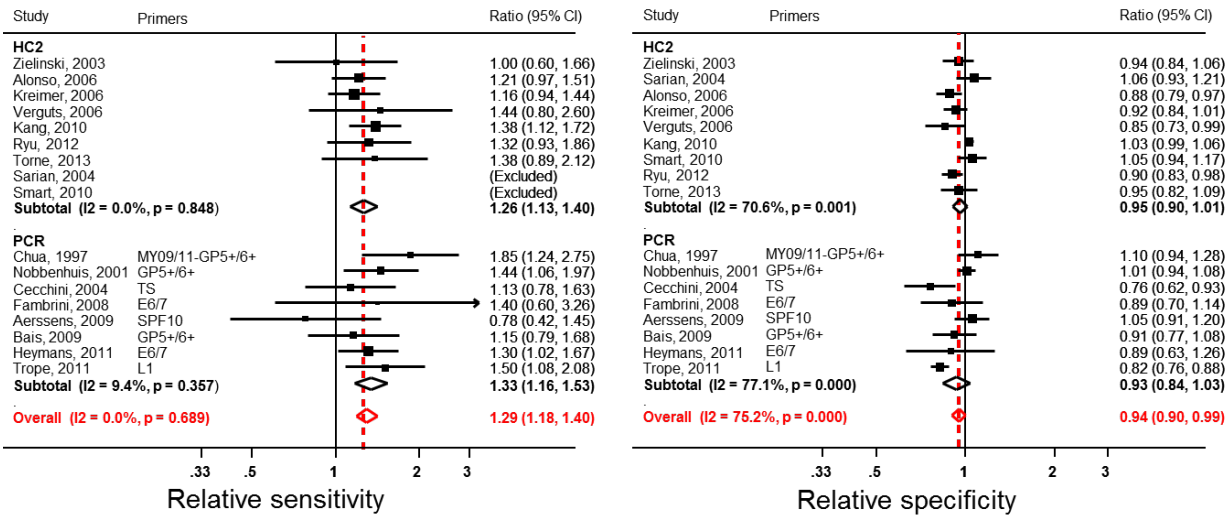


Figure 19. Study-specific and pooled relative sensitivity (left) and specificity (right) of hrHPV-DNA testing by HC2 and PCR versus cytology (cut-off: ASC-US), to predict occurrence of CIN2+ after treatment for CIN2+ or CIN3+.

9.7. Relative accuracy of combined hrHPV-DNA and cytology testing versus hrHPV-DNA testing alone to predict occurrence of CIN2+ after treatment for pre-cancer.

Combined testing for hrHPV-DNA and cytology at 3-9 months after treatment was equally sensitive (ratio 1.06, 95% CI: 0.97-1.17) but less specific (ratio 0.93, 95% CI: 0.89-0.97), compared to HPV-testing alone (Figure 20).

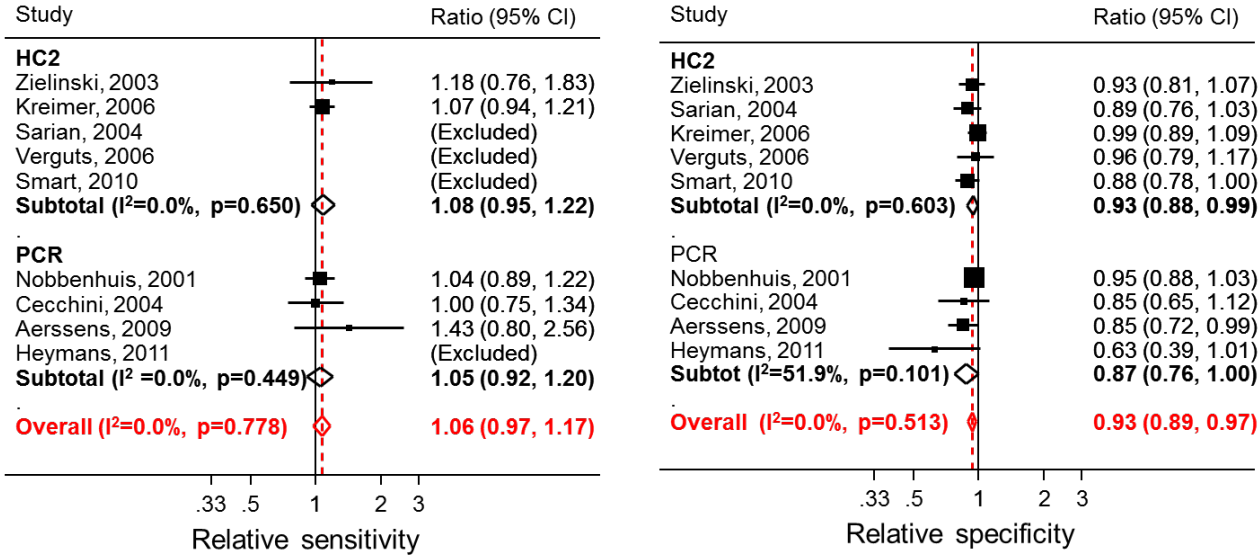


Figure 20. Study-specific and pooled relative sensitivity (left) and specificity (right) of combined testing with an hrHPV-DNA assay and cytology versus hrHPV-DNA testing alone, to predict occurrence of CIN2+ after treatment for CIN2+ or CIN3+.

9.8. Age as a risk factor for occurrence of disease after treatment

It has been suggested that higher age at time of treatment might be a risk factor for residual or recurrent disease after conservative treatment for high-grade CIN. However, conclusions are heterogeneous. In this report, a review of findings in literature is presented and discussed. Because different age cut-offs and definitions of disease were used in different studies and because study characteristics were too diverse, in general, meta-analytical pooling could not be performed.

A link between age and risk of disease after treatment is supported by the study of Tropé and colleagues, who demonstrated that the mean age of women who had recurrent CIN2+ (n=22) after treatment was significantly higher than women who remained disease free during 18 months of follow-up (43.2y vs. 37.2y, p<0.001)[153] (Table 14). Similarly, Verguts et al. reported a significant age difference among women and without residual/recurrent CIN2+ during 24 months of follow-up after treatment (51.5y vs. 39.8y, p=0.007)[161]. However, it must be noted that in other studies, although a higher mean age in the group of women with residual or recurrent disease was observed, the difference was not significant[97,162,169].
Six studies were identified that contained data on the occurrence of disease after treatment, stratified by age[97,154,165,170-172]. However, only three of them provided information on recurrent

disease defined as CIN2+[154,165,170], whereas in the others treatment failure was defined as recurrent CIN1+[97,171,172]. The absolute numbers and relative risks are listed in Table 14.

Valuable data is presented in the study of Flannelly and colleagues, which included 271 cases of residual/recurrent CIN1+ and 152 cases of residual/recurrent CIN2+ during 35 months of follow-up, among 3,426 treated women. In this large study, the risk of post-treatment disease was almost four times higher in women of 50 years and older (relative risk: 3.78 [95% CI=2.91-4.90)] [171]. In the same way, Prato and colleagues showed a clear link between age and post-treatment risk of CIN in women older than 35 versus younger women (relative risk: 5.45 [95% CI=1.25-23.80]). On the other hand, other studies failed to confirm age as a significant risk factor for residual or recurrent disease.
Additional valuable data is presented in two long-term surveillance studies. In the study of Melnikow and colleagues, based on 3,013 cases of CIN2/3 among 37,142 women treated and up to 15 years of surveillance, the authors concluded that women of 40 years or older and those with more severe index disease had higher rates of CIN 2/3 after treatment[99]. A nation-wide Swedish population-based linkage study, comprising more than 3 million women years, demonstrated an increased risk of invasive cervical cancer in women treated for CIN3+ compared to the general population. Among women treated after the age of 50 years or older, the cumulative incidence of invasive cervical cancer was significantly higher [102,173]. Previous observations that link older age to a higher risk of viral persistence after treatment, further support the notion that age is a risk factor for therapeutic failure [174-176].

9.9. Tables: Follow-up of women treated for cervical precancer

Table 13. Evaluation of the quality of each included study according to the QUADAS-2 check list[129].

	Risk of Bias													
	Patient Selection			Screening Test		Reference Test			Flow & Timing					
Author, Year	**P1**	**P2**	**P3**	**T1**	**T2**	**R1**	**R2**	**R3**	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**
Chua,1997[156]	Y	Y	Y	Y	?	Y	Y	Y	Y	Y	?	Y	Y	Y
Nobbenhuis,2001[164]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	?	Y	Y	Y
Zielinski,2003[158]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cecchini, 2004[165]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sarian,2004[159]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Alonso,2006[97]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y
Kreimer,2006[160]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Verguts,2006[161]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fambrini,2008[177]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Aerssens,2009[175]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Bais,2009[168]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Kang,2010[162]	Y	Y	Y	Y	N	Y	Y	Y	?	Y	?	Y	Y	Y
Smart,2010[163]	Y	Y	?	Y	N	Y	Y	Y	Y	Y	?	Y	Y	Y
Heymans,2011[157]	Y	Y	Y	Y	?	Y	Y	Y	Y		N	Y	Y	Y
Trope,2011[153]	Y	Y	?	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	
Ryu,2012[154]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Torne,2013[155]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	?	Y	Y	Y

QUADAS items: (P1) acceptable enrolment method, (P2) inappropriate exclusions avoided, (P3) acceptable follow-up time, (T1) pre-specified test cut-off, (T2) results of index and comparator tests blinded towards each other and reference test, (R1) acceptable reference test, (R2) results of reference test blinded towards index and comparator tests, (R3) incorporation bias avoided, (F1) acceptable delay between triage tests and reference test, (F2) partial verification avoided, (F3) differential verification avoided, (F4) withdrawals explained, (F5) non-interpretable results reported for tests, (F6) non-interpretable results reported for reference test. Each quality item is judged with: Y (fulfilled, green), ? (unclear, yellow), N (not fulfilled, red).

Table 14. Relative risk of occurrence of CIN after treatment for pre-cancer in older vs younger women.

Study	**Definition of disease**	**Age cut-off**	**Recurrence young age group (%)**	**Recurrence older age group (%)**	**RR (95% CI)**
Flannelly 2001[171]	CIN1+	≥50	215/3205 (6.7)	56/221 (25.3)	3.78 (2.91-4.90)
Cecchini 2004[165]	CIN2+	>39	8/67 (11.9)	2/17 (11.8)	0.99 (0.23-4.22)
Alonso 2006[97]	CIN1+	>45	26/165 (15.8)	10/38 (26.3)	1.67 (0.88-3.16)
Prato 2008[172]	CIN1+	>35	2/60 (3.3)	10/55 (18.2)	5.45 (1.25-23.80)
Ang 2011[170]	CIN2+	>35	41/1145 (3.6)	16/413 (3.9)	1.08 (0.61-1.91)
Ryu 2012[154]	CIN2+	≥50	11/165 (6.7)	1/18 (5.6)	0.83 (0.11-6.09)

10. Primary cervical cancer screening

10.1. Clinical accuracy of HPV tests to detect CIN2+ and CIN3+

10.1.1. Retrieved studies

The updated meta-analysis of the cross-sectional accuracy of HPV tests contains data from 60 studies, among which 9 randomised trials. In the large majority of studies, the HC2 or GP5+6+ PCR were used. The last couple of years data have been published on the performance of HPV-based screening with other HPV DNA or RNA detection systems (Cervista, Cobas-4800, *care*HPV, Papillocheck, APTIMA, Pretect HPV Proofer and others)[178].

10.1.2. Absolute cross-sectional accuracy of HPV testing

Hybrid Capture II

Overall, the sensitivity of HC2 (at standard conditions: RLU≥1) for detecting underlying CIN2+ and CIN3+ was 91% (95% CI: 89-93%) and 95% (95% CI: 94-97%), respectively (Figure 21, Figure 22 and Table 15).

The accuracy estimates were stratified geographically by developed and developing countries, separating China as a particular area. This stratification accounts for gold stand misclassification which was frequently notified in studies conducted in developing countries, where verification of the outcome was influenced by visual inspection[179,180]. The inter-study variation was substantially reduced in studies conducted in industrialised countries and was even non-significant in Chinese studies, where an improved gold standard was applied including colposcopy by experienced gynaecologists and random biopsy taking[181] (Figure 21 and Figure 22).

In European and North-American studies, the pooled sensitivity for CIN2+ was 96% (95% CI: 95-98%), whereas the pooled specificity was 91% (95% CI: 89-91%). The accuracy values of HC2 for CIN3+ were similar to those for CIN2+. Eleven percent (95% CI: 9-12%) of the screened population was hrHPV-positive.

Other HPV assays

The sensitivity of other HPV assays evaluated in primary screening was consistently high for detection of CIN3+ (in general >94%, except *care*HPV=87%). The specificity of these other tests varied between 86% and 91% (Table 15).

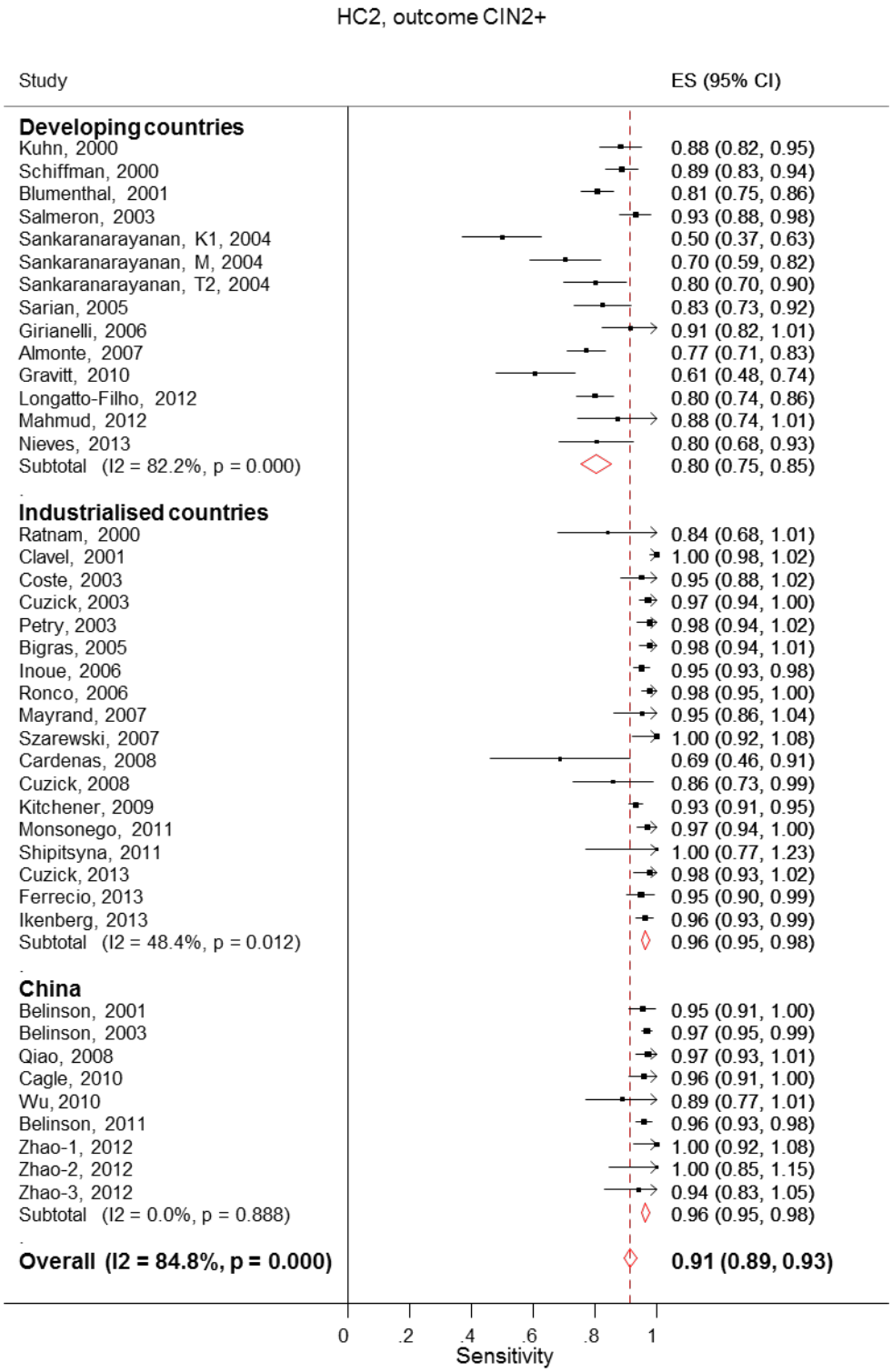


Figure 21. Meta-analysis of the sensitivity of HC2 as a primary screening test to detect CIN2+ in developing countries, industrialised countries and China.

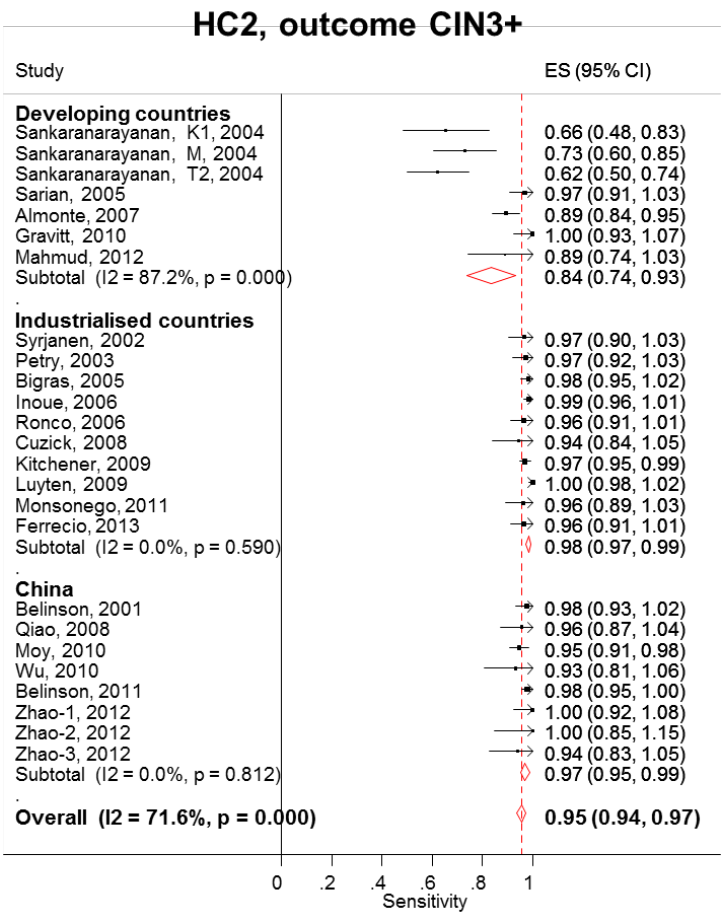


Figure 22. Meta-analysis of the sensitivity of HC2 as a primary screening test to detect CIN3+ in developing countries, industrialised countries and China.

10.1.3. Relative cross-sectional accuracy

The sensitivity of HC2 (at standard cut-off RLU≥1) was on average 24% to 43% higher than cytology, depending on the disease threshold (CIN2+ or CIN3+) or the cytological cut-off (ASC-US or LSIL) (Table 16). The specificity of HC2 for excluding CIN2+ was significantly lower than cytology: ratio of 0.98 (95% CI: 0.96-0.99) and 0.92 (95% CI: 0.90-0.94), considering the cut-offs ASC-US+ or LSIL+, respectively (Table 17).

The relative sensitivity and specificity of several HPV assays (GP5+/6+ PCR, *care*HPV, Cobas 4800, Abbott RT PCR, PapilloCheck, Cervista, BD HPV, Pretect HPV Proofer, APTIMA) compared to HC2 is displayed in Table 17. For the majority of tests, the 95% CIs around the accuracy measures always included unity with the exception of APTIMA, which was on average more specific than HC2 (ratio=1.07; [95% CI: 1.05-1.08]). Another exception was Pretect HPV Proofer which was also more specific (ratio=1.12; [95% CI: 1.10-1.13], based on one study) , but less sensitive than HC2 (ratio=0.74; [95% CI: 0.61-0.91]).

The combination of cytology with HC2 was, on average, 41% (95% CI: 36–47%) and 33% (95% CI: 29–37%) higher for the detection of CIN2+ or CIN 3+, respectively, than cytology alone (at

cut-off ASC-US+) (see Table 18). The specificity was 6% (95% CI: 6–7%) and 7% (95% CI: 7–8%) lower. In contrast, adding cytology to the HC2 increased the average sensitivity with only 7% (95% CI: 5–8%) for CIN2+ and 4% (95% CI: 3–6%) for CIN3+ compared to HC2 testing alone (Table 18). This resulted in a significant loss of specificity (ratios of 0.95; 95% CI: 0.94–0.96 and 0.94; 95% CI: 0.92–0.95), respectively.

10.1.4. Cross-sectional outcomes of randomised trials comparing cytology- with HPV based screening

No new studies could be identified than those included in a recent systematic review of the literature on three clinical applications of HPV testing, among which primary cervical cancer screening[69,182].Therefore, we shortly summarize the main findings of the aforementioned review.

Figure 23 displays the relative detection rate (which is equivalent to the relative sensitivity) for CIN2+ of hrHPV DNA testing in the experimental arm compared to cytology applied in the control arm. hrHPV testing consistently showed higher detection rates in the trials conducted in industrialised countries where conventional cytology was applied in the control arms. In one trial, conducted in India, comparing screening with cytology, VIA (visual inspection after application of acetic acid), hrHPV testing with HC2 assay and no screening[183], the relative detection rate (HC2 vs cytology) was lower than unity. These findings might be explained by a certain degree of gold-standard misclassification due to over-classifying non-high grade lesions which were HPV negative[179]. Evidence for these hypothesis was provided by observing a lower incidence of advanced cervical cancer and a lower mortality from cervical cancer in the HPV arm compared to the cytology arm[184]. Initial findings from the British trial, where LBC was used in the control arm, were interpreted as an indication that HC2 was not more sensitive than LBC[185] and that the consistent higher DR ration from the other RCTs conducted in Europe and North-America[186-190] were due to the use of conventional cytology in the control arm. Another explanation could be that some over-diagnosis occurred in LBC interpretation in the British trial. Evidence for this latter explanation is provided by the lower incidence of CIN3+ in the HPV arm of the UK trial (see below in sub-chapter 10.2.2) . If the DR ratio (which were not different from unity) were not based on bias, no decreased CIN3+ should be observed in the subsequent screening rounds among those who screened negative at baseline.

Similar findings were observed in the detection rate ratios for the outcome CIN3+ (see Figure 24).

In all trials, with an experimental arm with combined screening, adding cytology to hrHPV testing yielded only a minor and statistically insignificant increase in sensitivity (pooled DR ratios of 1.06 [95% CI: 0.97-1.16] and 1.04 [0.92-1.17] for CIN2+ and CIN3+, respectively), see Figure 25.

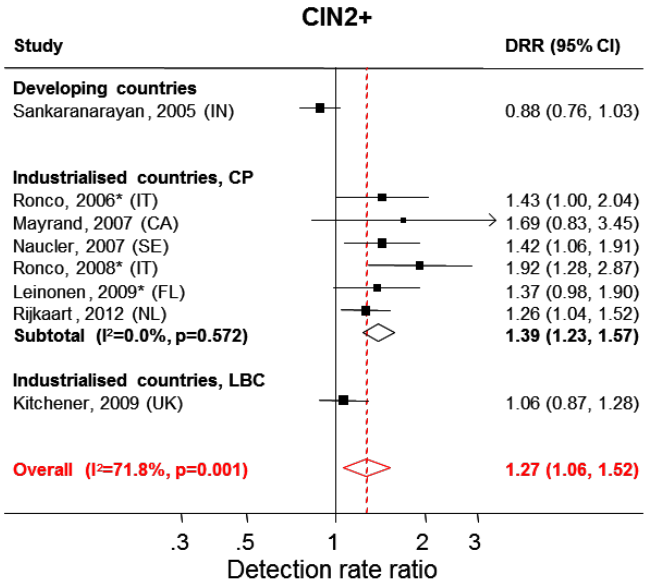


Figure 23. Meta-analysis of the detection rate ratio of CIN2+ in eight randomized trials identified by hrHPV testing versus cytology. Studies are grouped by type of cytology (conventional cytology [CP] or liquid-based cytology [LBC]) and by industrialized versus developing countries.
* restricted to women older than 35 years.

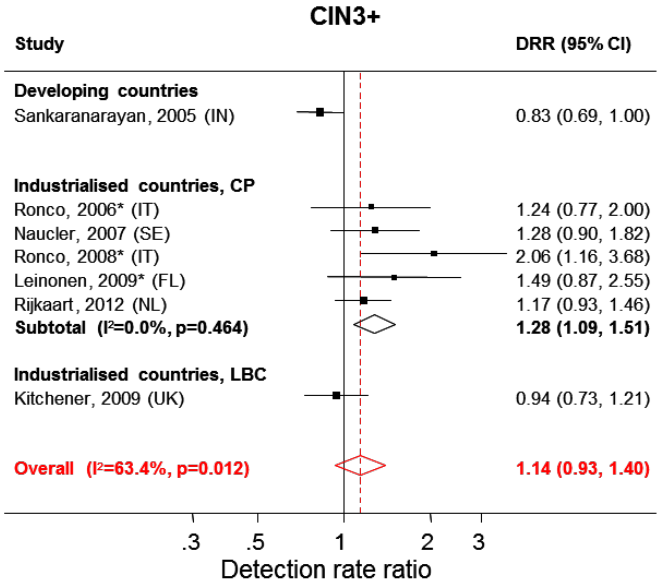


Figure 24. Meta-analysis of the detection rate ratio of CIN2+ in eight randomized trials identified by hrHPV testing versus cytology. Studies are grouped by type of cytology (conventional cytology [CP] or liquid-based cytology [LBC]) and by industrialized versus developing countries.
* restricted to women older than 35 years.

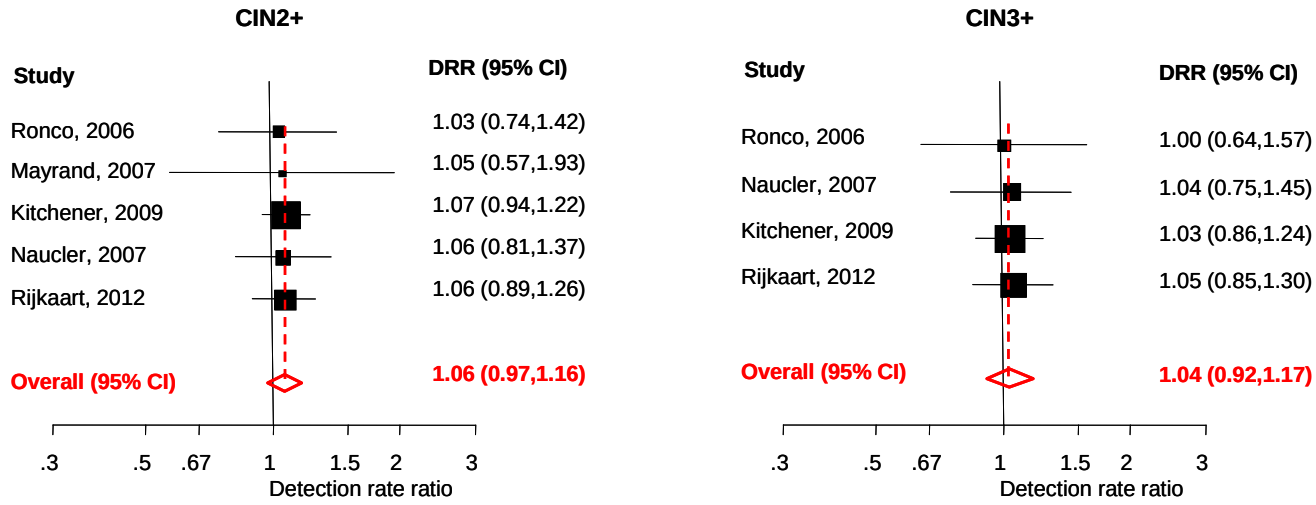


Figure 25. Relative sensitivity of HPV testing only vs combined testing (HPV + cytology).

10.2. Efficacy and effectiveness of HPV vs cytology in primary screening for cervical cancer

10.2.1. Retrieved studies

No new longitudinal results (second screening round) from other randomized trials could be identified besides the reports of four randomized trials conducted in Europe, respectively in Sweden, England, Italy and the Netherlands [191-194], which were already included in a previous review prepared by the Unit of Cancer Epidemiology[105]. No longitudinal outcomes, observed in subsequent rounds, were reported up hereto from the randomized trials conducted in Finland[195-197] or Canada (FOVAL trial)[198]. Another Canadian trial was excluded from our review since the criterion for the comparison groups was not fulfilled: randomization was based on the order of testing (collection for cytology or HPV as first or second test on the same individual)[187]. Moreover, no outcomes observed over subsequent screening rounds were foreseen[187].

10.2.2. Reduction in the cumulative incidence of CIN3+ and of cervical cancer

Figure 26 shows the detection rate ratio of CIN3 or worse lesions detected in the second screening round among women who were HPV negative (in the HPV arm), over those who were cytology negative (in the control arm). A very consistent reduction in the detection of CIN3+ is observed in all four trials (p for heterogeneity: 0.681).

Figure 27 shows the same statistic but now for invasive cervical cancer for the three trials that provided published data. Each of the trials were designed to demonstrate a reduction in CIN3+. Nevertheless, the ratios observed in two individual trials and also the pooled ratios were significantly lower than unity, indicating better protection against invasive cervical cancer associated with HPV-based compared to cytology-based screening.

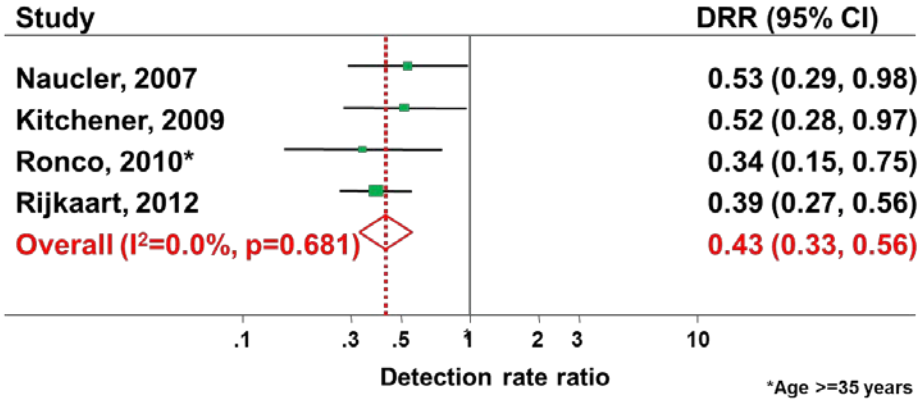


Figure 26. Meta-analysis of the main outcomes from randomised trials comparing HPV- and cytology-based cervical cancer screening. Relative detection rate of CIN3+, observed in the second screening round among women who were HPV-negative versus cytology-negative at enrolment. Source: Arbyn et al, *Vaccine* 2012[69].
* restricted to women of 35 years or older.

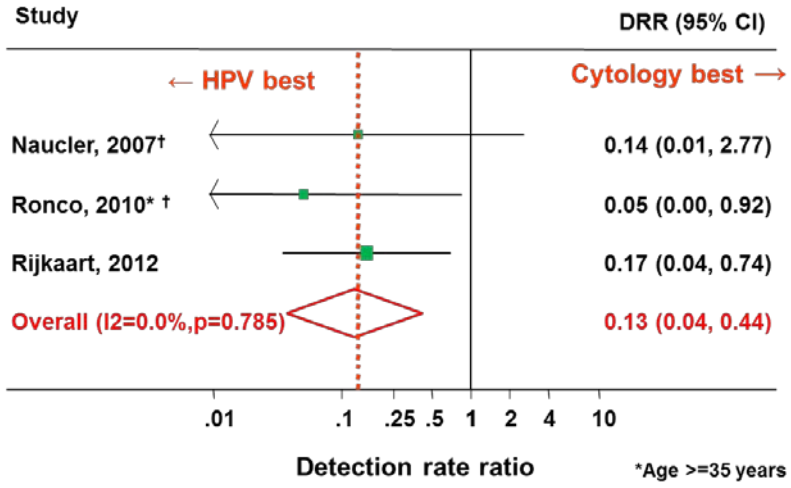


Figure 27. Meta-analysis of the main outcomes from randomised trials comparing HPV- and cytology-based cervical cancer screening. Relative detection rate of invasive cervical cancer, observed in the second screening round among women who were HPV-negative versus cytology-negative at enrolment.* restricted to women of 35 years or older; † continuity correction (+.5 in each cell because of zero cancer cases among HPV-negative women). Source: Arbyn, *Vaccine* 2012[69].

In addition, a recent pooled analysis of the individual data, from the four mentioned trials was published[199]. This pooled analysis provided more details regarding the protection against invasive cervical cancer by HPV-based compared to cytology-based screening (see Figure 28 and Figure 29), such as:

- the protective effect was observed only 2.5 years after screening (relative protection of 0·45; 95% CI: 0·25–0·81 versus 0·79, 95% CI: 0.46–1.36) before 2.5 years), but increased with follow-up time;
- the protective effect was similar for early (stage 1A) or advanced (stages ≥ 1A) cervical cancer;
- the protective effect was observed both in the total screened group (relative protection of 0·60; 95% CI: 0·40–0·89; see Figure 28) and in women with a negative screening test at base-line (relative protection of 0·30; 95% CI: 0.15–0.60; see Figure 29);
- there was no protective effect observed in the age group of <30 years (relative protection of 0·98 (95% CI: 0.19–5.20) (see Table 28);
- HPV-based screening protects more against adenocarcinoma (relative protection of 0.31, 95% CI: 0.14–0.69) than against squamous cancer (relative protection of 0.78, 95% CI: 0.49–1.25).

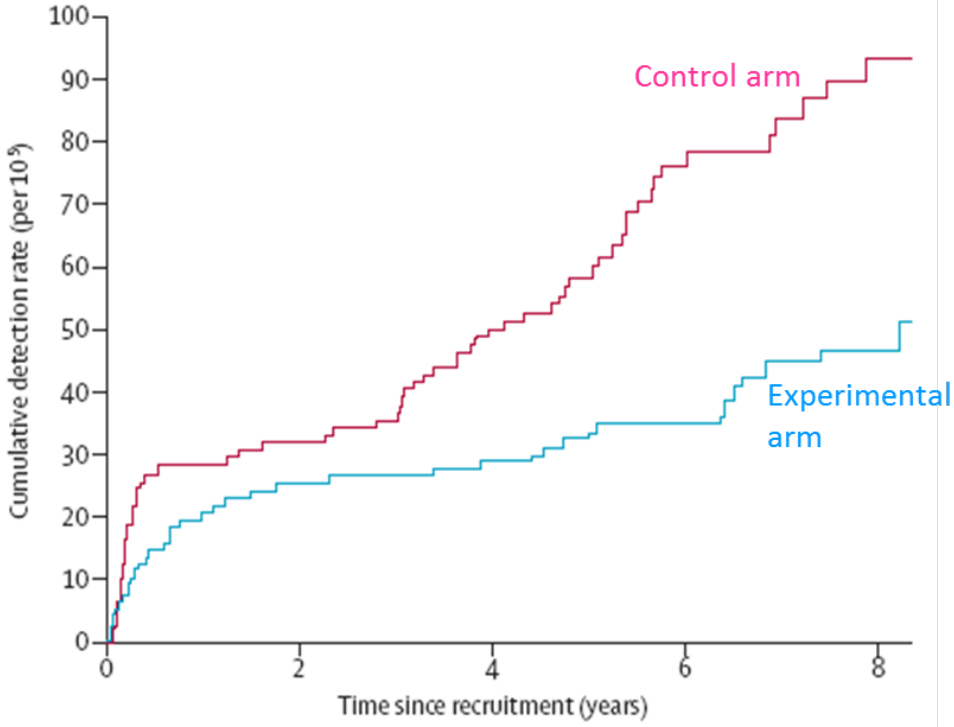


Figure 28. Cumulative detection rate of invasive cervical cancer among women included in the experimental arm (blue curve) screened with a hrHPV DNA test versus those included in the control arm (red curve) screened with cytology, irrespective of the screen test result (Ronco, Lancet 2013)[199].

Figure 29. Cumulative detection rate of invasive cervical cancer after a negative hrHPV DNA test (blue curve) compared to after a negative Pap test (red curve) (Ronco et al, Lancet 2014)[199].

10.3. Triage of HPV positive women

10.3.1. Absolute accuracy of triage algorithms

Retrieved studies

One American (ATHENA[200]) and six European trials (NTCC[186,201-203], ARTISTIC[185], SWEDESCREEN[204], VUSA[205], POBASCAM[206], and PUBLIC HEALTH TRIAL FINLAND[207]) contributed accuracy data for diverse triage strategies to manage women who had a positive hrHPV-DNA test at primary screening.

Short term cross-sectional outcome data were extracted for NTCC, SWEDESCREEN, ATHENA, and PUBLIC HEALTH TRIAL FINLAND. Longer term follow-up data were extracted for NTCC, ARTISTIC, VUSA and POBASCAM, comprising two to four years of follow-up. Five studies[186,200-203] had a complete design, referring all hrHPV-positive women to verification with the golden standard, while in the other five studies an incomplete design was applied, which means that only triage positive women were submitted to the golden standard.

Eight studies contained the absolute numbers of true- and false-positive and negative results for one-step triage with reflex cytology at cut-off ASC-US+[185,186,200,201,204-207]. Two reports from the Italian NTCC-1 trial[186,201] were combined for women between 35-60y and for women below 35y, respectively. Other triage strategies were less often described and could therefore not be meta-analysed.

Triage with reflex cytology (cut-off ASC-US+)

The pooled sensitivity and specificity for reflex cytology at cut-off ASC-US+ to detect CIN2+ was 79% (95% CI: 65-91%) and 79% (95% CI: 73-85%), respectively (upper part of Figure 30). The pooled sensitivity for CIN3+ was 82.0% (95% CI: 66.9-93.4%).

Addition of a second triage test for those having a first negative reflex cytology

Accuracy data for the addition of a second triage step after six months to manage women who had a negative cytology triage test at baseline, were available only from the POBASCAM trial[206] (lower part of Figure 30, Table 19). In this trial, the sensitivity of cytological triage at ASC-US+ for CIN3+ at baseline was 82% and by adding a second triage, the sensitivity increased to 96%, 100% and 100%, with the 2nd triage test being ASC-US+ cytology, hrHPV testing, or ASC-US+ cytology with hrHPV DNA testing. At the same time, the specificity decreased from 72% to 57%, 30%, or 28%. The PPV decreased from 35% to 23%, the NPV increased from 97.5% to 100% and the referral rate increased from 31% to 75%. Considering the outcome CIN2+, the sensitivity increased from 75% (baseline ASC-US+ triage) to 93%, 99.5% or 99.7%, whereas the specificity decreased from 88% to 80%, 44% or 41%, by adding one of the considered 2nd triage tests.

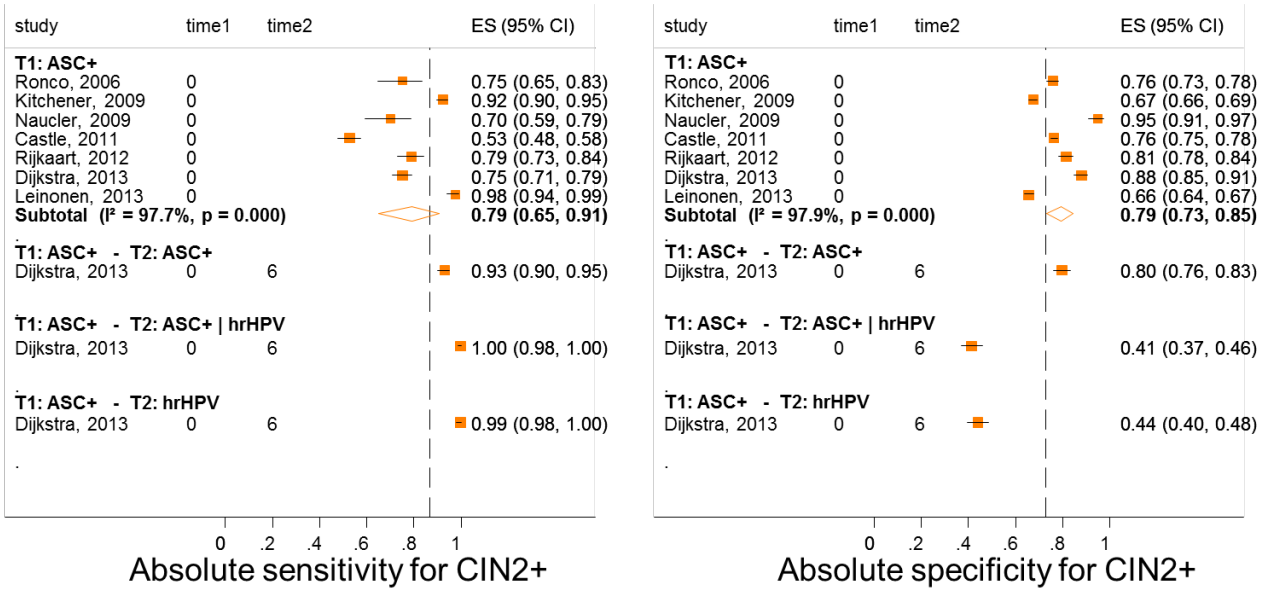


Figure 30. Meta-analysis of the absolute sensitivity and specificity to detect CIN2+ of four triage algorithms with reflex cytology (cut-off: ASC-US) as first triage. Time1 and time2 correspond to the timing of the triage step (in months).

Triage with a combination of reflex cytology (cut-off ASC-US+) and HPV16 or HPV1618 genotyping

From the POBASCAM and ATHENA trials, data could be extracted on reflex triage with a combination of cytology (cut-off ASC-US+) and HPV16 or HPV1618 genotyping, considering either using an 'OR' approach (one or both tests positive)[200,206] or an 'AND' approach (both tests positive)[200]. In Table 20, the absolute accuracy measures for the different triage algorithms are listed.

Based on these two studies, the pooled sensitivity and specificity of reflex triage with cytology (ASC-US+) or HPV1618 genotyping to detect CIN2+ was 86% (95% CI: 62-99%) and 62% (95% CI: 55-69%), respectively. To detect CIN3+, the pooled sensitivity and specificity were 90% (95% CI: 64.5-100%) and 53% (95% CI: 43-63%), respectively. Adding a second triage step using cytology (ASC-US+) or a combination of cytology and hrHPV-DNA testing resulted in a 13-14% gain in sensitivity for CIN2+ (98.7% [95%CI=96.9-99.4%] and 100% [95% CI=99-100%, respectively], and a ~2% gain in NPV (98.2% [95% CI=95.9-99.2%] and 100% [95% CI=97-100%], respectively). The referral rate increased up to 67% (95% CI= 64-70%) and 83% (95% CI=80-85%) when the second triage step was cytology or combined cytology-hrHPV respectively. Given the large contrast between ATHENA and POBASCAM, an intra-study (POBASCAM) comparison is indicated. In POBASCAM, reflex triage with ASC-US+ cytology and HPV1618 genotyping reached a sensitivity for CIN2+ of 94% (95% CI: 91-96%). Adding a second triage step using cytology (ASC-US+) or a combination of cytology and hrHPV-DNA testing resulted in a 5-6% gain in sensitivity for CIN2+ (reaching 99% and 100%, respectively), and a 5-6% gain in NPV (up to 98% and 100%, respectively). The referral rate increased up to 67% and 83% when the second triage step was cytology or combined cytology-hrHPV respectively. The accuracy of reflex ASC-US+ combined with HPV1618 genotyping reached a high sensitivity for CIN3+ (97%), adding a 2nd triage step resulted in a sensitivity of 99-100% but was accompanied by a specificity loss of 18-37%.
In a triage algorithm where both cytology (ASC-US+) and the HPV1618 genotyping test had to be positive, the sensitivity dropped considerably (down to 30% [95% CI=26-35%] for CIN2+, and 34%[95% CI=29-40%] for CIN3+), while the specificity to exclude CIN2+ increased (up to 93% [95% CI=92-94%].
The ATHENA trial contained data on accuracy of triage with a combination of cytology (ASC-US+) and HPV16 genotyping or HPV1618[200]. Genotyping for HPV16 or HPV1618 showed a sensitivity for CIN3+ of 75% (95%=CI 69-80%) and 78% (95% CI=73-83%), respectively, whereas the specificity to exclude CIN2+, was 64% (95% CI=63-66%) and 59% (95% CI=57-61%), respectively.

Triage with reflex cytology (cut-off LSIL+) and combinations with HPV16 or HPV1618 genotyping.

In the lower part of Table 20, the absolute accuracy measures are listed for the different triage algorithms with reflex cytology at cut-off LSIL. Four studies were identified which provided accuracy estimates for reflex triage with LSIL+ cytology[192,201,207,208], whereas for reflex LSIL+ triage combined with genotyping for HPV16 or HPV1618, only the ATHENA trial provided accuracy data.
The pooled sensitivity and specificity of reflex cytology at cut-off LSIL+ to detect CIN2+ was 68.4% (95% CI: 41.5-90.0%) and 86.8% (95% CI: 83.4-89.8%), respectively. The pooled sensitivity to detect CIN3+ was 70.4% (95% CI: 42.9-91.7%).

Addition of HPV1618 genotyping to reflex cytology at cut-off LSIL+, resulted in only a small gain in sensitivity for CIN3+ (from 70% to 72%) but resulted in a substantial drop in specificity (from 84% to 65%) and doubled the referral rate (from 18% to 38%). In contrast, if both tests had to be positive, the referral rate dropped considerably (to 6%), but also significantly lowered the sensitivity for CIN3+ (to 27%) and the NPV for CIN3+ to 94%.

Given the high variability among the four studies that contributed data for cytology triage at cut-off LSI+, an intra-study (ATHENA) comparison might be more appropriate to appreciate the change in accuracy for adding HPV1618 genotyping to reflex cytology at cut-off LSIL+. This intra-study comparison shows a substantial gain in sensitivity (+29% for CIN2+, +32% for CIN3+) and a drop in specificity (-21% for both CIN2+ and CIN3+).
Addition of genotyping for HPV1618 to genotyping for HPV16 only, yielded a 4% sensitivity gain, but a 7% loss in specificity.

10.3.2. Relative accuracy of cytology and/or hrHPV-DNA based triage algorithms versus reflex cytology (ASC-US+)

The relative accuracy of the different triage algorithms were compared with reflex cytology at cut-off ASC-US+.
Data on the triage with combined cytology (ASC-US+) or HPV1618 genotyping, versus cytology (ASC-US+) alone was available in the POBASCAM and ATHENA trial [200,206] (Figure 31 and Figure 32). When either cytology or the HPV1618 genotyping test had to be positive, a significantly higher sensitivity (ratio: 1.32 [95% CI=1.16-1.51] for CIN2+) but lower specificity (ratio:0.77 [95% CI=0.74-0.79] for CIN2+) was observed, compared to cytology testing alone. Results were similar for CIN3+.
Four studies allowed comparison of reflex cytology triage at cut-of LSIL+ versus ASC-US+[185,186,200,201,207]. Using LSIL+ as cut-off resulted in a 16% drop in sensitivity (ratio: 0.84 [95% CI=0.74-0.95] for CIN2+), but a 22% increase in specificity (ratio: 1.22 [95% CI=1.12-1.33] for CIN2+).

For most triage algorithms, data were available for only one study (see Belgian Knowledge Centre-238 Report, for a detailed overview[111]). Recalling women with normal reflex cytology for a second triage test with cytology at cut-off ASC-US+ after 6 months (T1:ASC-US+, T2:ASC-US+), resulted in a significantly increased sensitivity (ratio: 1.24 [95% CI= 1.16-1.32] for CIN2+), but a drop in specificity (ratio: 0.91 [95% CI= 0.86-0.96] for CIN2+) in the study of Dijkstra et al[206]. Using a hrHPV-DNA assay in the second triage step further increased sensitivity, but specificity was halved compared to reflex cytology (ASC-US+) alone[111].

The largest gain in sensitivity was observed when reflex cytology was combined with HPV16 or HPV1618 genotyping and only one of both assays had to be positive (sensitivity ratio: 1.35 [95% CI=1.20-1.51] and 1.41 [95% CI=1.26-1.58], respectively)[200]. However, this was linked with a significant drop in specificity (ratio: 0.84 [95% CI=0.82-0.87] and 0.77 [95% CI=0.74-0.80], respectively).
Reflex triage with HPV1618 genotyping was as sensitive (ratio: 0.99 [95% CI=0.86-1.13] for CIN2+) and as specific (ratio: 0.99 [95% CI=0.96-1.02] for CIN2+) as reflex cytology at cut-off ASC-US. Reflex triage with HPV16 genotyping was less sensitive (ratio: 0.84 [95% CI=0.72-0.97] for CIN2+) but more specific (ratio: 1.09 [95% CI=1.06-1.12] for CIN2+) compared to reflex cytology.
Triage algorithms using a higher cytology cut-off (LSIL+ or HSIL+) or where both cytology and a genotyping test had to be positive resulted in significantly lower sensitivities compared to reflex cytology.

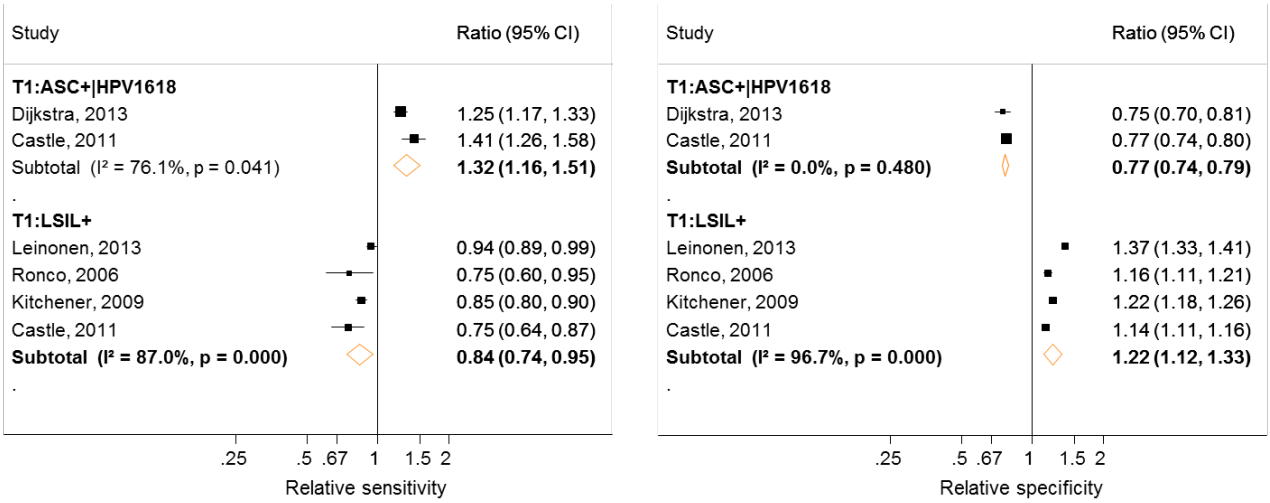


Figure 31. Relative sensitivity (left) and specificity (right) of two scenarios compared to reflex cytology at cut-off ASC-US+ to detect CIN2+ in women with a positive hrHPV DNA screening test (restricted to scenarios where a pooling from at least 2 studies was possible).

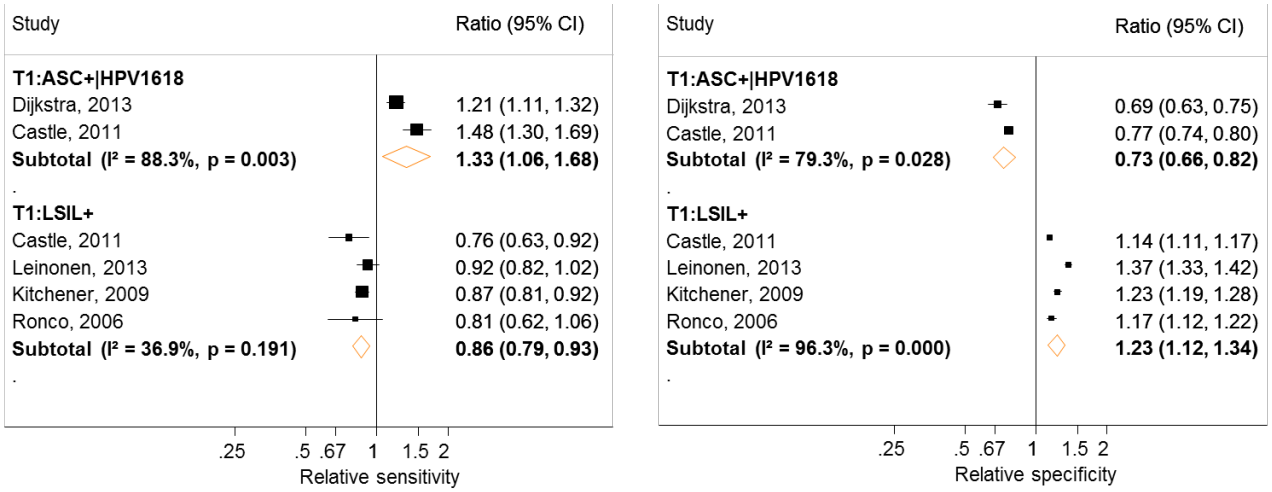


Figure 32. Relative sensitivity (left) and specificity (right) of two scenarios compared to reflex cytology at cut-off ASC-US+ to detect CIN3+ in women with a positive hrHPV DNA screening test. (restricted to scenarios where a pooling from at least 2 studies was possible).

10.3.3. p16^{INK4a} immuno-cytochemistry

Phase-2 of the Italian NTCC trial contained data on the use of a molecular biomarker (overexpression of p16^{INK4a}) to triage hrHPV-positive women. Two reports provided the cross-sectional[202] and longitudinal[203] outcomes of triage based on p16-immunocyto-chemistry**.** The cross-sectional sensitivity and specificity of reflex triage with p16^{INK4a}, was 88.0% (95%

CI=79.6-93.9) and 60.6% (95% CI=57.5-63.6%), respectively, for CIN2+, and 90.5% (95% CI= 77.4-97.3%) and 58.4% (95% CI= 55.5-61.4%), respectively, for CIN3+[202].
The longitudinal sensitivity for $p16^{INK4a}$ assessed over three years was lower than the cross-sectional sensitivity regarding detection of CIN2+ (88.0% vs. 79.0%) and CIN3+ (90.5% vs 84.6%), indicating disease development in women with a negative reflex $p16^{INK4a}$ test [209].
By lack of an intra-study comparator, no relative accuracy could be derived for p16-based triage in NTCC-2. However, by comparing with absolute accuracy measures for ASC-US+ triage from NTCC-1 (cross-sectional sensitivity of 75.3% [95% CI: 64.5-84.2] for CIN2+, 82.1% [95% CI: 66.5-92.5%] for CIN3+; specificity of 75.8% [95% CI: 73.1-78.3%] for excluding CIN2+, we may also obtain credible relative accuracy estimates. Allowing for this inter-NTCC comparison, we can conclude that p16-based triage was 1.17 times (95% CI: 1.01-1.35) more sensitive to detect CIN2+ and 1.10 times (95% CI: 0.92-1.32) to detect CIN3+ compared to simple cytology triage at cut-off ASC-US. The specificity of p16 triage to exclude CIN2+ was significantly lower than ASC-US+ triage: ratio of 0.80 (95% CI: 0.75-0.85).

Using $p16^{INK4a}$/Ki-67 detection to manage hrHPV-positive/cytology-negative women
Another study was identified containing data on triagewith $p16^{INK4a}$/Ki-67 double staining of women who were hrHPV-positive but had normal cytology at primary screening[210].
In this study, $p16^{INK4a}$/Ki-67 triage had a specificity to exclude CIN2+ of 80.9% (95% CI= 76.7-84.7) and a sensitivity of 91.9% (95% CI=78.1-98.3) for CIN2+, and 96.4% (95% CI=81.7-99.9) for CIN3+.

10.4. Screening interval

The screening interval should be defined taking into account the cumulative risk of significant disease after a previous negative screening test. Other elements that should be considered are: the attendance at future screening, available resources and the burden of other health problems.

Figure 33 shows the cumulative risk to develop CIN3+ in the next five years according to the initial combined HPV and cytology status (pooled from 5 European [211-216] and two American screening cohorts)[217,218]. The risk was very low for HPV-negative women (0.2% and 1.2%, for women without or with cytological abnormalities, respectively). However, for women who were hrHPV positive, this risk was substantially higher: 6% or 12%, for women without or with ASC-US or worse cytology at baseline, respectively. Follow-up of the European cohorts showed that the low risk of CIN3+ associated with a negative HPV test extended for at least six years (Figure 34) [211]. The Portland study also provided outcomes regarding the risk of developing CIN3+ over 10 years in a population screened by cytology only where masked hrHPV testing was performed at baseline: 10.2% (if ASC-US+), 6.9% (if hrHPV+), 6.8% (if hrHPV+ or ASC-US+), 1.4% (if cyto-negative), 0.9% (if hrHPV-negative) and 0.8% (if negative for both cytology and hrHPV) [218] More recently, follow-up outcomes over a period of 18 years were published, confirming the low risk of CIN3+ (<1%) up to 18 years after a negative hrHPV DNA result (0.79%; 95 % CI: 0.63% to 0.98%)[219].

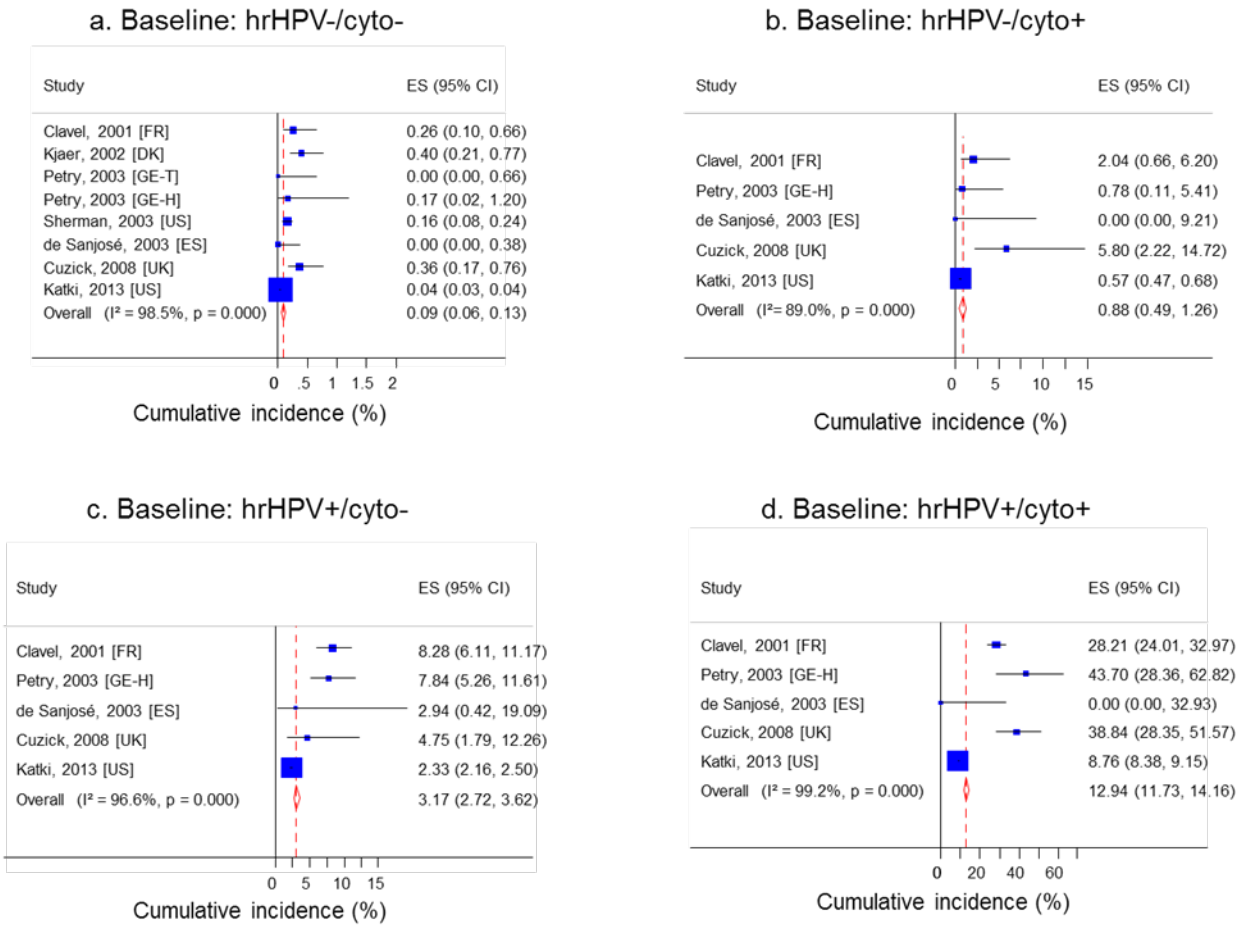


Figure 33. Five-year cumulative incidence of CIN3+ among women attending cervical cancer screening who are at baseline cyto-negative/hrHPV-negative (a), cyto-positive/hrHPV-negative (b), cyto-negative/hrHPV-positive (c) and cyto-positive/hrHPV-positive (d). Pooled from 6 European and 2 American screening cohorts. * The Portland study reported cumulative incidences over ~4 years [218]. Updated from Arbyn et al, *Vaccine* 2012[69].

In the Kaiser-Permanente cohort, where women of 30 years or older were screened by cytology and HC2, the 5 year cumulative incidence of invasive cervical cancer was: 0.90% if positive for both hrHPV and cytology, 0.54% if hrHPV positive and cytology-negative, 0.16% if hrHPV-negative and cytology positive and 0.016% if negative for both (Figure 35)[220]. The 5-year cervical cancer risk corresponding with a negative hrHPV DNA result was 0.19% indicating that negative cytology stratified slightly the low-risk associated with one negative HPV test.

The recent pooled analysis of the individual data of four European randomised trials comparing cytology- with HPV-based screening showed a significantly lower risk after a negative hrHPV DNA test compared to a after negative cytology result[199]. The lower detection rate was only observable after two and a halve years and remained observable up to eight years (see Figure 29).

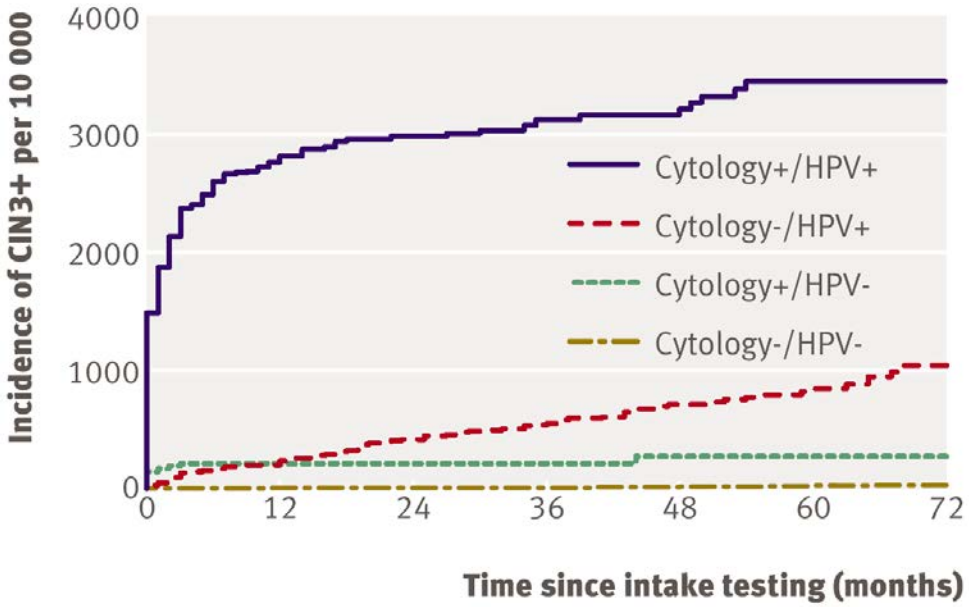


Figure 34. Cumulative incidence of CIN3 or cervical cancer in five European cohorts screened at baseline with cytology and hrHPV DNA testing, according to the joint cytology/HPV test results at enrolment. Source: Dillner et al, *BMJ* 2008[211].

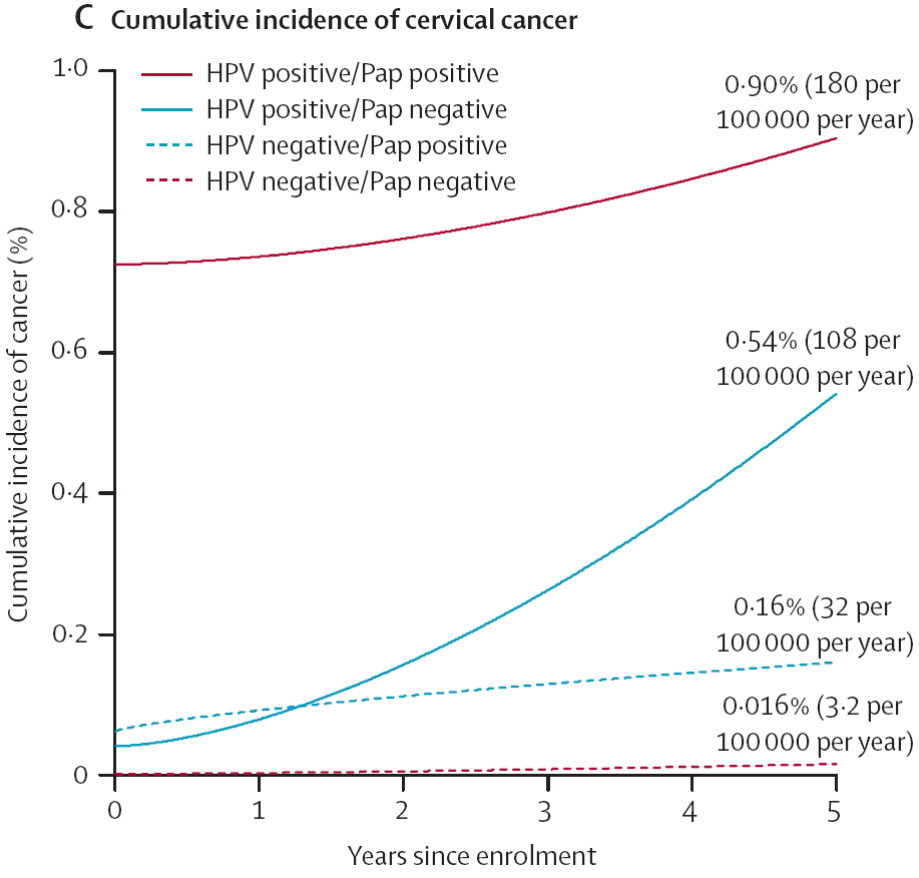


Figure 35. Cumulative incidence of cervical cancer according the joint HPV/cytology status at baseline (Katki et al, *Lancet Oncol* 2011)[220].

10.5. Target age group for HPV-based screening

10.5.1. Principles

The second edition of the European guidelines for cytology-based screening propose starting screening in the age range 20–30, but preferentially not before age 25 years. Screening should be continued until the age of 60 or 65 years depending on the burden of disease and available resources [221-223]. It is recommended to provide screening at 3- to 5-year intervals for these target age groups[221,224]. Stopping screening in older women is probably appropriate among women who have had three or more consecutive previous (recent) normal cytology results.

Screening should be initiated a few years before the age where the incidence of cervical cancer starts to rise[225].

An additional condition to determine the target age group is that efficacy or effectiveness of screening is demonstrated within the age range concerned. A recent large case-control study conducted in the UK, has demonstrated that cytological screening at age <25 was not effective (odds ratio for cervical cancer among screened vs non-screened not < 1)[226].

The age-specific incidence of cervical cancer in Belgium is shown in Figure 36. The incidence of invasive cervical cancer in Belgium starts rising in the age groups 25-34. We can therefore conclude that the Belgian screening policy (one Pap smear every three years for women aged 25-64 years)[227] is in agreement with the principle outlined above and in line with the European recommendations.

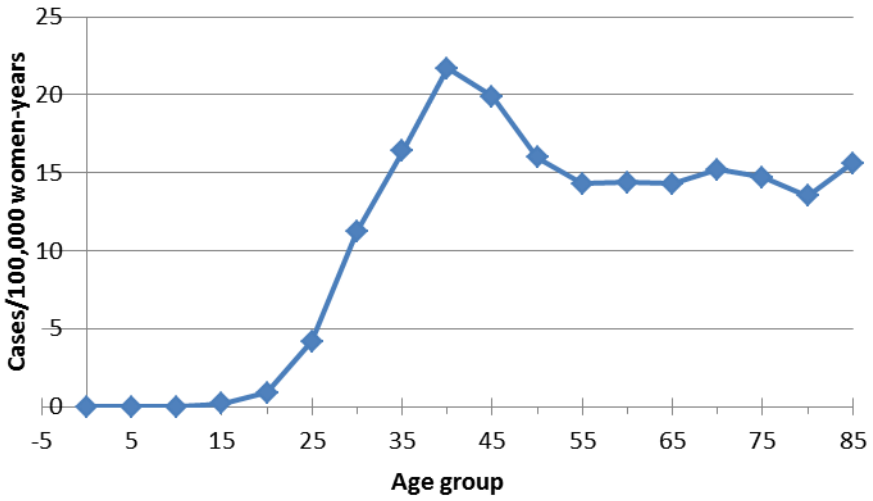


Figure 36. Age-specific incidence of cervical cancer in Belgium (average for the years 2008-2010). (Source: Belgian Cancer Registry).

10.5.2. Age-specific prevalence of HPV infection

In developed countries, the prevalence of HPV infections peaks shortly after onset of sexual activity and typically peaks in older teenagers and women in their early 20ies. Thereafter, the prevalence decreases progressively by age with sometimes a discrete peak around 45-55 years[39,67,228].

Figure 37 shows the age-specific prevalence of high-risk HPV in Belgium[229]. The shapes of the curves for hrHPV as well as for individual hrHPV genotypes (data not shown) are as observed in most other West-European countries[67,229].

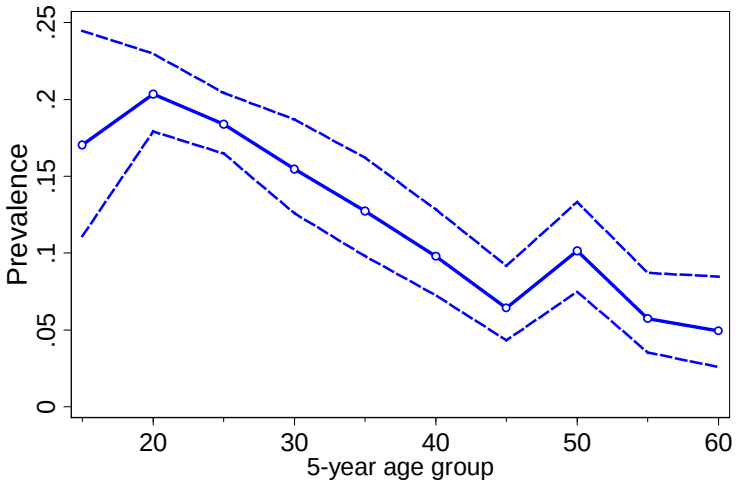


Figure 37. Prevalence of hrHPV in women attending cervical cancer screening in Belgium (2011-2014). Source: SEHIB study (Arbyn, et al, *Cancer Epidemiology* 2016)[144].

10.5.3. Variation in test accuracy of cervical cytology and hrHPV DNA testing by age.

Data included in a published pooled analysis of the age-specific accuracy of virological and cytological screening for cervical pre-cancer (CIN2+ or CIN3+) [230] derived from six European [212,214,231-233] and two North-American studies[234,235] were reanalysed using a random effect model for meta-analysis. Data were aggregated in three age groups: <30 years, 30-49 year, and 50 or older.

A summary of the pooled absolute and relative accuracy parameters for both CIN2+ and CIN3+ is presented in Table 21.

The sensitivity of hrHPV testing was homogenously high (>94%, p for heterogeneity by age group >0.10), whereas the sensitivity of cytology was low and increased significantly by age (<52% for women aged <50 years, ~80% for women of older age).

In all age-groups, the sensitivity of hrHPV DNA testing was significantly and substantially higher than that of cytology, but the contrasts decreased by age. The inter-study variability in cytology was substantially greater than that of hrHPV testing, reflecting the lower reproducibility of the former test.

The pooled specificity of hrHPV DNA testing and cytology increased by age: between 88% and 94%, for the former, and between 96% and 98% for the latter. The specificity of hrHPV testing always was significantly lower than that of cytology, but the contrasts decreased by age: ratio of 0.91 among women younger than 35, and 0.96 among women older than 50.

The higher relative sensitivity of hrHPV and lower specificity compared to cytology in young women may reflect the transient character of HPV infection and a certain degree of over-diagnosis[236].

Over-diagnosis was also suggested from the 2nd phase of the Italian NTCC screening trial. In this phase all women testing positive with the HC2 randomized to the experimental arm. The detection rate of CIN3+ was more than three times higher than in the cytology arm (ratio=3.1, 95%CI 2.2-4.4) among women aged 25-34. In women aged 35-60, the relative detection rate was substantially lower (ratio= 1.68, 95% CI 1.25-2.26) [237].

10.5.4. Influence of age on screening efficacy

Pooling of the individual data of four randomised clinical trials comparing HPV- with cytology-based screening, conducted in Europe, allowed addressing the effect modification by age group on screening efficacy (reduction in the incidence of invasive cervical cancer (Table 22). HPV screening below the age of 30 was not more protective against cancer than cytology.

10.5.5. Age to stop screening

When HPV testing is used, screening could be stopped at a younger age, given the longer lead time compared to cytology based screening and the lower risk of acquiring new progressive infections after a negative screening test[238]. However, in the supplements to the 2nd edition of the European guidelines, the lack of data is recognized regarding the future risk of cervical cancer after a previous negative HPV test among women older than forty[110,239]. For safety reasons, no change in screening policy with respect to stopping screening could be proposed[110].

10.5.6. Age-specific adverse effects related to treatment of screen-detected lesions

Data from recent meta-analyses show higher rates of adverse pregnancy outcomes in women with a prior history of excisional treatment of cervical pre-cancer than in the general population[89,90,240], in particular when the excision is deep or a large proportion of cervical tissue is excised [241]. Adverse pregnancy outcomes may be caused by cervical incompetence or decreased protection against ascending infections and include preterm premature rupture of the membranes, preterm delivery (<37 weeks of gestation) and low birth weight (<2500 gr). Shallow excision of the transformation zone might be free of adverse obstetrical effects as suggested from recent reports where excisions probably were less aggressive [242-244].
The increased risk of adverse obstetric outcomes should be an issue when defining start of screening, in particular when HPV-based screening is considered[245].

10.6. Validation of HPV assays for primary cervical cancer screening

Randomised trials have demonstrated that HPV-based screening using the HC2 assay or the GP5+/6+ PCR with EIA identification of 14 high-risk HPV types, is more effective in reducing the incidence of cervical cancer than high-quality cytology-based screening. Therefore these assays are considered as clinically validated[105,246]. The question raises whether new HPV assays require assessment of efficacy through such trials. An international team of experts have defined cross-sectional clinical equivalency criteria (Meijer *et al, Int J Cancer*, 2009) allowing claims for

other HPV DNA assays for use in primary screening[247]. It is accepted that a new assay detecting HPV DNA with at least similar accuracy as HC2 or the GP5+/6+-EIA PCR would be as efficacious in cervical cancer screening. For clinical applications of HPV testing it is crucial to define detection of a relevant cervical pre-cancer lesion as an outcome. Analytical sensitivity, which addresses the capacity of a test to pick very low copy numbers of HPV DNA, is considered as not clinically meaningful since it does not allow to distinguish between clinically irrelevant (regressing) from relevant (progressing) infections[248]. Therefore it is not sufficient that new assays just fulfil analytical accuracy prequalification criteria. Clinical validation of a new assay must be a necessary condition before acceptance in clinical practice.

The candidate test should demonstrate non-inferior sensitivity and specificity for CIN2+ compared to HC2 or GP5+/6+ PCR (test for non-inferiority of matched proportion <0.05, accepting the benchmarks of ≥0.90 for relative sensitivity and ≥0.98 for relative specificity)[249]. A representative set of samples (minimally 60 CIN2+ cases, 800 ≤CIN1 cases) derived from a population-based screening cohort should be selected[247]. Moreover, a high reproducibility (lower confidence bound ≥ 87%) should be reached.
A systematic search of published peer-reviewed references was performed from PUBMED and Embase, completed with citations of the Meijer guideline in www. scopus.com[250].

Data regarding eight hrHPV DNA tests (Abbott RealTime High Risk HPV test [Abbott, Wiesbaden, Germany], BD Onclarity HPV assay [BD Diagnostics, Sparks, MD, USA], Cervista HPV HR Test [Hologic, Bedford, MA, USA], cobas 4800 HPV test Roche Molecular System [Pleasanton, CF, USA], type-specific qPCR targeting E6/E7 genes [in house, AML, Antwerp, Belgium], HPV-Risk assay [Self-Screen BV, Amsterdam, The Netherlands], PapilloCheck HPV-screening test [Greiner Bio-One, Frickenhausen, Germany] and Anyplex II HPV HR [Seegene, Seoul, Korea]) and one hrHPV RNA assay (APTIMA HPV assay [Hologic, Bedford, MA, USA]) were retrieved, where the clinical performance for detecting high-grade CIN was assessed according to the Meijer guideline[247]. Two other newer HPV DNA assays were assessed within the rigorous VALGENT validation framework[251]: LMNX genotyping kit HPV GP (Diassay B.V., Rijkswijk, the Netherlands) and Xpert HPV (Cepheid, Sunnyvale, USA). For two other HPV DNA tests (*care*HPV Test [Qiagen, Gaithersburg, MD, USA] and MALDI-TOF [Medical Microbiology and Clinical Chemistry, Lund University, Malmö University Hospital, Sweden]) and a five-type HPV RNA test (Pretect HPV-Proofer [NORCHIP, Klokkarstua, Norway]), comparative accuracy data from cross-sectional screening studies were retrieved[250]. The characteristics of these tests are summarised in Table 23. The outcomes of the validation assessment regarding non-inferiority with respect to sensitivity and specificity for CIN2+ and reproducibility, are summarised in Table 24 and Table 25, respectively.

The cobas 4800 HPV Test, the Abbott RealTime High Risk HPV test, the BD Onclarity HPV assay, the PapilloCheck HPV-screening test, and the Anyplex II HPV HR were consistently validated in two or three studies, whereas the HPV-Risk assay and the Xpert HPV were validated each in one study. Other tests which partially fulfil the 2009-guidelines are the following: Cervista HPV HR Test, PCR-LMNX, an in-house *E6/E7* RT quantitative PCR and MALDI-TOF (matrix-assisted laser desorption-ionization time-of-flight). The APTIMA HPV assay targeting *E6/E7* mRNA of hrHPV was also fully validated. However, the cross-sectional equivalency criteria of the 2009-guidelines were set up for HPV DNA assays. Demonstration of a low risk of

CIN3+ after a negative APTIMA test over a period of 5 years or longer is waited for to inform about its utility and safety in cervical cancer screening.

As indicated in Table 15 and Table 17, several other hrHPV DNA tests may be sufficiently accurate but are not yet validated using the criteria of Meijer et al. [150]. Besides accuracy, also other characteristics will have to be taken in account in future choices of tests, such as high throughput capacity, costs, applicability on samples taken by the woman (see Chapter 11) and the possibility to perform ancillary triage tests.

10.7. Tables: Primary screening

Table 15. Meta-analyses on the test performance of HPV DNA testing in primary screening, using different HPV DNA assays to detect CIN2+ or CIN3+.

HPV assay	Test cut-off	Outcome	Studies	Pooled sensitivity (95% CI)	Range	Pooled specificity (95% CI)	Range	Pooled test-positivity rate (95% CI)
HC2	RLU≥1	CIN2+	41	0.91 (0.89-0.93)	0.50-1.00	0.89 (0.87-0.90)	0.61-1.00	0.13 (0.12-0.14)
		CIN2+	18°	0.96 (0.95-0.98)	0.69-1.00	0.91 (0.89-0.92)	0.85-0.96	0.11 (0.09-0.12)
		CIN3+	26	0.95 (0.94-0.97)	0.62-1.00	0.89 (0.88-0.91)	0.82-0.95	
		CIN3+	10°	0.98 (0.97-0.99)	0.94-1.00	0.91 (0.88-0.93)	0.85-0.95	
GP5+/6+	+ signal	CIN2+	3	0.95 (0.94-0.97)	0.94-1.00	0.95 (0.93-0.96)	0.86-0.96	0.08 (0.05-0.10)
Cervista	+ signal	CIN3+	1	0.95 (0.92-0.99)	-	0.90 (0.90-0.91)	-	0.11 (0.10-0.12)
Abbott	+ signal	CIN3+	1	0.95 (0.85-1.00)	-	0.87 (0.86-0.88)	-	0.13 (0.13-0.14)
Aptima	+ signal	CIN3+	4	0.99 (0.94-1.00)	0.96-1.00	0.91 (0.90-0.93)	0.90-0.93	0.09 (0.08-0.11)
Cobas 4800	+ signal	CIN3+	2	0.95 (0.87-1.00)	0.92-1.00	0.88 (0.81-0.96)	0.84-0.92	0.12 (0.05-0.20)
*care*HPV	RLU≥1	CIN3+	1	0.87 (0.73-1.00)	-	0.86 (0.85-0.88)	-	0.15 (0.13-0.16)
MALDI-TOF	+ signal	CIN3+	1	0.94 (0.91-0.98)	-	0.89 (0.89-0.90)	-	0.12 (0.11-0.13)

° Restricted to studies conducted in North-America or Europe.

Table 16. Relative accuracy of virologic versus cytologic screening to find CIN2+ or CIN3+, using HC2 to detect high-risk HPV and conventional or liquid systems to detect cytological abnormalities.

Comparison	Outcome	Relative sensitivity	Range	Relative specificity	Range	No. of studies*
HC2 / Cytology (ASC-US+)	CIN2+	1.24 (1.17, 1.32)	0.91-2.93	0.98 (0.96, 0.99)	0.86-1.11	38/33
HC2 / Cytology (LSIL+)	CIN2+	1.38 (1.27, 1.51)	0.89-2.73	0.92 (0.90, 0.94)	0.67-1.03	26/23
HC2 / Cytology (ASC-US+)	CIN3+	1.32 (1.15, 1.51)	0.96-2.63	0.98 (0.97, 1.00)	0.88-1.11	26/22
HC2 / Cytology (LSIL+)	CIN3+	1.40 (1.23, 1.60)	0.97-2.52	0.93 (0.91, 0.95)	0.84-1.03	18/15

+ 1st number is the number of studies included for sensitivity/2nd number is the number of studies for specificity.

Table 17. Relative accuracy of other HPV tests compared to HC2 (at RLU≥1) to find underlying CIN2+ or CIN3+in primary screening.

Comparison	**Outcome**	**Relative sensitivity** (95% CI)	**Relative specificity** (95% CI)	**No. of studies**
GP5+6+ / HC2	CIN2+	1.00 (0.96-1.04)	0.99 (0.91-1.07)	2
*care*HPV / HC2	CIN2+	0.93 (0.85-1.01)	0.98 (0.96-1.01)	1
Cobas 4800 / HC2	CIN2+	1.00 (0.96-1.04)	0.99 (0.98-1.00)	3
Abbott / HC2	CIN2+	1.00 (0.96-1.03)	1.02 (1.01-1.03)	3
Papillocheck / HC2	CIN2+	0.99 (0.96-1.04)	0.99 (0.98-1.00)	1
BD HPV / HC2	CIN2+	1.00 (0.93-1.07)	0.99 (0.97-1.00)	1
Pretect HPV Proofer / HC2	CIN2+	0.74 (0.61-0.91)	1.12 (1.10-1.13)	1
APTIMA / HC2	CIN2+	0.98 (0.94-1.03)	1.04 (1.01-1.08)	5

Table 18. Relative accuracy of combined testing with HC2 (at RLU≥1) and cytology (ASC-US+), compared to HC2 or cytology, to detect CIN2+ or CIN3+ in primary screening.

Comparison	**Outcome**	**Relative sensitivity (95% CI)**	**Relative specificity (95% CI)**	**No. of studies***
HC2-CYTO(ASC-US+)/ CYTO(ASC-US+)	CIN2+	1.41 (1.36-1.47)	0.94 (0.93-0.94)	14
	CIN3+	1.33 (1.29-1.37)	0.93 (0.92-0.93)	10/9
HC2-CYTO(ASC-US+) / HC2	CIN2+	1.07 (1.05-1.08)	0.95 (0.94-0.96)	14/11
	CIN3+	1.04 (1.03-1.06)	0.94 (0.92-0.95)	10/7

* 1st number is the number of studies included for sensitivity/2nd number is the number of studies for specificity.

Table 19. Absolute sensitivity, specificity, positive and negative predictive values, and referral rate for triage with reflex cytology at cut-off ASC-US+ combined or not with second triage step among women with a positive hrHPV test.

Triage1	Triage2	No studies	Outcome	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Referral rate % (95% CI)
ASC-US+		**7**[6]	CIN3+	82.0 (66.9-93.4)	72.3 (67.0-77.3)	22.9 (13.7-33.6)	97.5 (95.0-99.1)	32.0 (27.0-37.3)
ASC-US+	ASC-US+	1	CIN3+	95.9 (91.4-98.1)	57.2 (53.5-60.8)	32.2 (28.0-36.7)	98.5 (96.8-99.3)	52.1 (48.8-55.5)
ASC-US+	hrHPV	1	CIN3+	100 (97.5-100)	30.2 (26.9-33.7)	23.3 (20.2-26.8)	100 (98.2-100)	75.1 (72.0-77.9)
ASC-US+	ASC-US+ | hrHPV	1	CIN3+	100 (97.5-100)	28.2 (25.0-31.7)	22.8 (19.7-26.2)	100 (98.1-100)	76.7 (73.7-79.4)
ASC-US+		7	CIN2+	79.5 (65.2-90.8)	79.1 (73.0-84.6)	44.8 (27.2-63.1)	94.3 (89.4-97.8)	32.0 (27.0-37.3)
ASC-US+	ASC-US+	1	CIN2+	93.0 (89.9-95.2)	79.8 (75.9-83.2)	78.2 (74.1-81.8)	93.6 (90.8-95.6)	52.1 (48.8-55.5)
ASC-US+	hrHPV	1	CIN2+	99.5 (98.1-99.9)	44.0 (39.6-48.5)	58.1 (54.2-61.9)	99.1 (96.6-99.7)	75.1 (72.0-77.9)
ASC-US+	ASC-US+ | hrHPV	1	CIN2+	99.7 (98.5- 100)	41.3 (36.9-45.7)	57.0 (53.2-60.8)	99.5 (97.2-99.9)	76.7 (73.7-79.4)

[66] For comparison in Dijkstra et al 2013: the accuracy measure of triage with ASC-US+ cytology for CIN3+ SE=80%, SP=69%, PPV=35%, NPV=94%, referral rate=31%; and for CIN2+ SE=75%, SP=88%, PPV=83%, NPV=82%.

Table 20. Absolute sensitivity, specificity, positive and negative predictive values, and referral rate of reflex cytology (using ASC-US+ [upper part] or LSIL [lower part] as cut-off) combined with HPV16 or HPV1618 genotyping to triage women with a positive hrHPV test.

Triage1	Triage2	Number of studies	Outcome	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Referral rate % (95% CI)
ASC-US+ | HPV1618		2[7]	CIN3+	89.6 (64.5-100)	52.8 (42.6-62.8)	19.7 (6.9-37.0)	97.9 (96.1-99.2)	52.6 (37.3-67.5)
ASC-US+ &HPV1618		1	CIN3+	34.1 (28.6-40.2)	92.3 (91.3-93.1)	25.5 (21.2-30.4)	94.8 (93.9-95.5)	9.6 (8.7-10.6)
ASC-US+ | HPV1618	ASC-US+	1	CIN3+	99.3 (96.3-99.9)	39.8 (36.3-43.5)	25.9 (22.5-29.7)	99.6 (98.0-99.9)	67.0 (63.8-70.1)
ASC-US+ | HPV1618	ASC-US+|hrHPV	1	CIN3+	100 (97.5-100)	20.8 (17.9-23.9)	21.1 (18.3-24.3)	100 (97.4-100)	82.9 (80.2-85.3)
ASC-US+ | HPV16		1	CIN3+	74.6 (68.9-79.6)	63.3 (61.7-65.0)	13.6 (11.9-15.5)	97.0 (96.2-97.6)	39.4 (37.8-41.0)
ASC-US+ & HPV16		1	CIN3+	28.6 (23.3-34.4)	94.4 (93.6-95.2)	28.5 (23.3-34.3)	94.5 (93.6-95.2)	7.2 (6.4-8.1)
ASC-US+ | HPV1618		2[3]	CIN2+	85.7 (61.9-98.9)	62.1 (55.1-68.9)	42.0 (3.2-88.8)	94.6 (93.2-95.9)	52.6 (37.3-67.5)
ASC-US+ &HPV1618		1	CIN2+	30.0 (25.6-34.8)	92.9 (91.9-93.7)	33.8 (29.0-39.0)	91.6 (90.6-92.5)	9.6 (8.7-10.6)
ASC-US+ | HPV1618	ASC-US+	1	CIN2+	98.7 (96.9-99.4)	57.7 (53.2-62.0)	64.6 (60.5-68.4)	98.2 (95.9-99.2)	67.0 (63.8-70.1)
ASC-US+ | HPV1618	ASC-US+|hrHPV	1	CIN2+	100 (99.0-100)	30.5 (26.6-34.8)	52.9 (49.2-56.6)	100 (97.4-100)	82.9 (80.2-85.3)
ASC-US+ | HPV16		1	CIN2+	71.1 (66.3-75.4)	64.4 (62.7-66.1)	19.6 (17.6-21.7)	94.8 (93.8-95.7)	39.4 (37.8-41.0)
ASC-US+ & HPV16		1	CIN2+	25.8 (21.6-30.4)	95.0 (94.2-95.7)	38.7 (32.9-44.9)	91.3 (90.3-92.2)	7.2 (6.4-8.1)
LSIL+		**4[8]**	**CIN3+**	**70.4 (42.9-91.7)**	**84.3 (80.1-88.1)**	**17.7 (12.9-23.1)**	**98.2 (95.5-99.7)**	**18.2 (13.1-23.9)**
LSIL+ | HPV1618		1	CIN3+	72.2 (66.4-77.4)	65.2 (63.5-66.8)	13.9 (12.1-15.8)	96.8 (96.0-97.5)	37.5 (35.9-39.1)
LSIL+ & HPV1618		1	CIN3+	27.4 (22.2-33.2)	95.5 (94.7-96.1)	31.9 (26.1-38.4)	94.4 (93.6-95.2)	6.2 (5.4-7.0)
LSIL+ | HPV16		1	CIN3+	68.3 (62.3-73.7)	71.8 (70.2-73.3)	15.8 (13.7-18.1)	96.7 (95.9-97.3)	31.1 (29.6-32.6)
LSIL+ & HPV16		1	CIN3+	22.2 (17.5-27.8)	96.6 (95.9-97.1)	33.3 (26.6-40.8)	94.1 (93.3-94.9)	4.8 (4.1-5.6)
LSIL+		**4[4]**	**CIN2+**	**68.4 (41.5-90.0)**	**86.8 (83.4-89.8)**	**33.4 (24.6-42.8)**	**96.4 (92.2-99.0)**	**18.2 (13.1-23.9)**
LSIL+ | HPV1618		1	CIN2+	67.9 (63.0-72.4)	66.2 (64.5-67.8)	19.6 (17.6-21.9)	94.4 (93.4-95.3)	37.5 (35.9-39.1)
LSIL+ & HPV1618		1	CIN2+	23.2 (19.2-27.7)	95.9 (95.1-96.5)	40.7 (34.4-47.4)	91.1 (90.1-92.0)	6.2 (5.4-7.0)
LSIL+ | HPV16		1	CIN2+	64.2 (59.3-68.9)	72.9 (71.3-74.5)	22.4 (20.0-25.0)	94.4 (93.4-95.2)	31.1 (29.6-32.6)
LSIL+ & HPV16		1	CIN2+	19.2 (15.6-23.5)	97.0 (96.3-97.5)	43.5 (36.2-51.0)	90.8 (89.8-91.7)	4.8 (4.1-5.6)

[7] In Dijkstra 2013, the performance parameters of triage with reflex testing with ASC-US+ cytology and HPV1616 testing for CIN2+ were: SE=94.1%, SP=65.9%, PPV=68.3%, NPV=93.4% and referral rate=60.4%; and for CIN3+: SE=97.3%, SP=47.4%, PPV=28.2%, NPV=98.8%

[8] In Castle 2011, the performance parameters of triage with reflex testing with LSIL+ cytology for CIN2+ were: SE=39.2%, SP=86.7%, PPV=26.4%, NPV=92.1% and referral rate=16.1%; and for CIN3+: SE=40.1%, SP=85.8%, PPV=17.9%, NPV=94.9%

Table 21. Sensitivity and specificity of hrHPV DNA testing and cytology (at ASC-US+) to detect CIN2+ and CIN3+ in women attending cervical cancer screening, by age group. Relative sensitivity and specificity of the two tests (adapted from Cuzick, IJC 2006)[230].

Absolute accuracy of hrHPV DNA testing

		<35	35-49	50+	trend
Sensitivity	CIN2+	98.4 (96.3-100)	95.2 (91.4-99.1)	99.3 (92.7-100.0)	0.13
	CIN3+	98.8 (96.5-100)	94.5 (89.5-99.0)	100.0 (93.2-100.0)	0.86
Specificity	CIN2+	88.4 (84.9-92.0)	93.5 (91.5-95.5)	94.4 (91.8-97.1)	<0.001
	CIN3+	86.4 (82.6-90.3)	93.1 (91.0-95.2)	94.1 (91.4-96.9)	<0.001

Absolute accuracy of cytology at ASC-US+

		<35	35-49	50+	trend
Sensitivity	CIN2+	46.4 (25.9-66.8)	50.5 (34.9-66.1)	80.1 (67.5-92.7)	<0.001
	CIN3+	49.4 (29.8-69.0)	49.0 (33.3-64.6)	80.8 (67.5-94.1)	0.01
Specificity	CIN2+	95.7 (9379-97.7)	97.3 (96.2-98.4)	98.1 (96.9-99.24)	<0.001
	CIN3+	94.7 (92.2-97.1)	97.0 (95.8-98.2)	97.9 (96.6-99.2)	<0.001

Relative accuracy (hrHPV DNA vs. cytology)

		<35	35-49	50+
Sensitivity	CIN2+	2.10 (1.49-2.94)	1.80 (1.29-2.50)	1.19 (1.00-1.40)
ratio	CIN3+	1.91 (1.39-2.63)	1.83 (1.29-2.59)	1.22 (1.01-1.46)
Specificity	CIN2+	0.92 (0.90-0.95)	0.96 (0.94-0.98)	0.96 (0.93-0.99)
ratio	CIN3+	0.91 (0.89-0.94)	0.96 (0.94-0.98)	0.96 (0.93-0.99)

Table 22. Relative risk or protective effect (reduction in incidence of invasive cervical cancer) in women screened with HPV testing vs cytology, according to age at enrolment (Source: Ronco, Lancet 2013)[199].

Age at enrolment (years)	RR	95% CI	I^2	p for heterogeneity (inter-study)
<30	0·98	(0·19–5·20)	0.00%	0.34
30-34	0.36	(0·14–0·94)	7.20%	0.36
35-49	0.64	(0·37–1·10)	0.00%	0.55
>=50	0.68	(0·30–1·52)	36.50%	0.21

RR= relative risk (cumulative incidence of invasive cervical cancer observed in the HPV arm compared to that observed in the cytology arm). I^2= measure of inconsistency, the percentage of total variation across studies due to inter-study heterogeneity.

Table 23. Characteristics of hrHPV tests assessed for cross-sectional equivalence performance criteria used to validate them for cervical cancer screening.

	hrHPV assay	**Nucleic acid targeted**	**Type of amplification**	**Genes Targeted**	**Separate genotyping**	**Internal control for human genes**
Standard comparator tests						
*	HC2	DNA	Signal	Several (undefined)	No	No
*	GP5+/6+ PCR-EIA	DNA	Target	*L1*	No	No
HPV assays validated according to the Meijer protocol						
1	APTIMA HPV assay[252]	RNA	Target	*E6/E7*	No (a prototype identifying HPV16, 18-45 exists)	No. Assay includes internal controls for non-infectious RNA and DNA.
2	Abbott RealTime High Risk HPV test [253-255]	DNA	Target	*L1*	16,18 & 12 other hr types	*β-globin*
3	BD Onclarity HPV assay[256]	DNA	Target	*E6/E7*	16,18,31,45,51,52; 33-58; 56-59-66; 35-39-68	*β-globin*
4	Cervista HPV HR Test[257,258]	DNA	Signal	*L1/E6/E7*	14 hr types. Separate typing of 16, 18 available as a reflex test.	human *histone 2*
5	cobas 4800 HPV test[259,260]	DNA	Target	*L1*	16,18 & 12 other hr types (see HC2 plus 66)	*β-globin*
6	qPCR(E6/E7)[261]	DNA	Target	*E6/E7*	Separate typing of hrHPV types 16,18,31,33, 35,39,45,51,52,56,58,59,66,68; phr type 53 and lr types 6 & 11.	*β-globin*
7	HPV-Risk assay[262]	DNA	Target	*E7*	16, 18 and 13 other hr types (see HC2 plus 66 & 67)	*β-globin*
8	PapilloCheck HPV-screening test[263]	DNA	Target	*E1*	Separate typing of 16,18,31,33,35,39,45, 51,52,53,56,58,59,66,68. As well as phr types 70,73,82 and lr types 6,11,40,42,43,44.	*ADAT1*
9	Anyplex II HPV HR[264]	DNA	RT PCR	*L1*	No*	*β-globin*

	hrHPV assay	Nucleic acid targeted	Type of amplification	Genes Targeted	Separate genotyping	Internal control for human genes
HPV assays assessed in cross-sectional studies in a primary screening setting or according to the VALGENT protocol						
1	GP5+/6+-LMNX[265]	DNA	Target	*L1*	Separate typing of hr HPV types 16,18, 31,33,35,39,45,51,52,56,58,59,66,68; Detects also phr types:26,53, 73, 82.	Human DNA fragment located on chromosome 14
2	*care*HPV Test[266]	DNA	Signal	Several (undefined)	No	No
3	MALDI-TOF[267]	DNA	Target	*L1*	Separate typing of hr HPV types 16,18, 31,33,35,39,45,51,52,56,58,59,66,68.	*ß-globin*
4	Pretect HPV-Proofer[268]	RNA	Target	*E7/E7*	Yes (16,18,31,33, 45)	U1 small nuclear ribonucleo-protein-specific mRNA.
5	Xpert HPV[269]	DNA	RT PCR	*E6/E7*	Separate identification of HPV16, and four groups of types (HPV18/45, HPV31/33/52/58, HPV51/59, and HPV39/56/66/68).	*hydroxymethylbilane synthase (HMBS)*

hrHPV, high-risk human papillomavirus; *ADAT1* gene, adenosine deaminase tRNA specific 1; phr, potentially high-risk; qPCR, quantitative PCR; MALDI-TOF, matrix-assisted laser desorption-ionization time-of-flight.

***** HC2 and GP5+/6+ PCR-EIA are clinically validated in randomised efficacy trials and therefore used as standard comparator tests to validate other HPV assays.

****** The complementary assay Anyplex II HPV28 enables separate identification of 28 HPV types: 14 high-risk types (16,18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 66 and 68), 6 possibly high-risk types (HPV26, 53, 69, 70, 73, and 82) and 8 low-risk types (6, 11, 40, 42, 43, 44, 54, 61). CT: cycle threshold.

Table 24. Sensitivity and specificity of hrHPV assays validated for cervical cancer screening, relative sensitivity and specificity of the evaluated hrHPV assays compared to the standard comparator tests (HC2 or GP5+/6+ PCR-EIA).

		Evaluated assay		Comparator assay			Evaluated/comparator assay		Non-inferiority test*		Validation level‡
		Absolute			Absolute		Relative				
Evaluated assay	Study	sensitivity	specificity	Comparator assay	sensitivity	specificity	sensitivity	specificity	p_{sens}	p_{spec}	
GP5+/6+ EIA	Meijer, 2009[150]	98.7%	96.0%	HC2*	98.7%	94.1%	1.00	1.02	0.0037	<0.0001	⊕⊕⊕⊕
PapilloCheck	Hesselink, 2010[263]	95.8%	96.7%	GP5+/6+ EIA	96.4%	97.7%	0.99	0.99	<0.0001	0.0072	⊕⊕
	Heard, 2016 [270]	96.1%	89.7%	GP5+/6+ EIA	94.1%	90.4%	1.02	0.99	0.0002	0.0970	
Abbott RT hrHPV test	Carozzi, 2011[253]	96.4%	92.3%	HC2	97.6%	92.6%	0.99	1.00	0.0040	0.0087	⊕⊕⊕
	Poljak, 2011[254]	100.0%	93.3%	HC2	97.4%	91.8%	1.03	1.02	0.0112	0.0000	
	Hesselink, 2013[255]	95.6%	92.0%	GP5+/6+ EIA	98.5%	91.8%	0.97	1.00	0.0278	0.0003	
cobas 4800	Heideman, 2011[259]	90.0%	94.6%	HC2	91.7%	94.4%	0.98	1.00	0.0216	0.0009	⊕⊕⊕
	Lloveras, 2013[260]	98.3%	86.2%	HC2	98.3%	85.3%	1.00	1.01	0.0093	0.0012	
qPCR(*E6/E7*)	Depuydt, 2012[261]	93.5%	95.6%	HC2	83.9%	94.4%	1.11	1.01	0.0001	<0.0001	⊕⊕
APTIMA	Heideman, 2013[252]	95.5%	94.5%	GP5+/6+ EIA	100.0%	93.6%	0.96	1.01	0.0394	0.0002	x
Cervista	Boers, 2014[257]	89.0%	91.2%	HC2	93.4%	88.8%	0.95	1.03	0.0043	<0.0001	⊕
	Alameda, 2015[258]	98.4%	85.2%	HC2	100.0%	86.4%	0.98	0.99	0.0122	0.3170†	
BD Onclarity	Ejegod, 2014[256]	92.9%	87.7%	HC2	94.2%	88.8%	0.99	0.99	0.0009	0.0216	⊕⊕⊕
	Cuschieri, 2015 [271]	96.7%	89.6%	HC2	98.4%	89.9%	0.98	1.00	0.0245	0.0155	
	Ejegod, 2016 [272]	96.1%	89.7%	GP5+/6+ PCR	94.1%	90.4%	1.02	0.99	0.0002	0.0970	
HPV-Risk assay	Hesselink, 2014[262]	97.1%	94.3%	GP5+/6+ EIA	97.1%	94.1%	1.00	1.00	0.0056	0.0003	⊕⊕
Anyplex II HPV HR	Hesselink, 2016 [273]	98.3%	93.6%	GP5+/6+ PCR	98.3%	94.1%	1.00	0.99	0.0052	0.0232	⊕⊕⊕
	Jung, 2016 [274]	92.5%	81.7%	HC2	87.5%	81.8%	1.06	1.00	0.0067	0.0354	
Xpert HPV	Cuschieri, 2016 [275]	94.1%	90.3%	GP5+/6+ PCR	94.1%	90.3%	1.00	1.00	0.0171	0.0269	⊕⊕

* p values for non-inferiority of the evaluated assay compared to the comparator assay.

† We corrected an error in Alameda, 2015[258], which was due to switching of + and – columns and rows. The corrected data showed that the non-inferiority test was not significant for specificity.

‡Validation level for the test accuracy criterion as proposed by Meijer *et al* 2009[150].

⊕⊕⊕⊕ validated in large randomized controlled trials with cancer incidence as an outcome; considered as standard comparator tests;

⊕⊕⊕: fully validated in multiple studies; ⊕⊕ fully validated in one study; ⊕ partially validated.

X not evaluated since not a hrHPV DNA assay.

Table 25. Intra- and inter-laboratory reproducibility of hrHPV assays validated for cervical cancer screening.

Evaluated Assay	Study	Intra-laboratory reproducibility (hrHPV) Overall	LCIB	Kappa	Inter-laboratory reproducibility (hrHPV) Overall	LCIB	Kappa	Validation level‡
PapilloCheck	Hesselink, 2010[263]	97.6%	96.3%	0.941	94.0%	92.1%	0.842	⊕⊕
Abbott RT hrHPV test	Carozzi, 2011[253]	98.5%	97.3%	0.969	-	-	-	⊕⊕
	Poljak, 201[254]1	100.0%	99.5%	1.000	100.0%	99.5%	1.000	
	Hesselink, 2013[255]	99.8%	99.1%	0.996	98.4%	97.2%	0.965	
cobas 4800	Heideman, 2011[259]	98.3%	97.2%	0.963	94.6%	92.8%	0.882	⊕⊕
	Lloveras, 2013[260]	98.3%	97.2%	0.963	98.4%	97.2%	0.962	
qPCR(*E6/E7*)	Depuydt, 2012[261]	98.7%	97.8%	0.956	-	-	-	⊕
APTIMA	Heideman, 2013[252]	96.0%	94.4%	0.893	95.1%	93.3%	0.865	⊕⊕
Cervista	Boers, 2014[257]	91.9%	89.7%	0.829	90.7%	88.4%	0.807	⊕⊕
	Alameda, 2015[258]	94.9%	93.0%	0.890	96.5%	94.9%	0.907	
BD Onclarity	Ejegod, 2014[256]$	98.6%	97.5%	0.967	98.4%	97.2%	0.962	⊕⊕
	Ejegod 2016 [272]†	97.4%	95.9%	0.935	96.8%	95.2%	0.921	
HPV-Risk assay	Hesselink, 2014[262]	99.5%	97.3%	0.987	99.2%	97.2%	0.981	⊕⊕
Anyplex II HPV HR	Hesselink, 2016 [273]	96.0%	94.4%	0.909	96.8%	95.3%	0.927	⊕⊕
	Jung, 2016 [274]	98.0%	96.7%	0.980	97.4%	96.0%	0.974	
Xpert HPV	Cuschieri, 2016 [275]	96.9%	95.0%	0.924	97.8%	96.2%	0.948	⊕⊕

hrHPV: high-risk human papillomavirus; LCIB: lower 95% confidence interval bound; RT:real time; qPCR: quantitative polymerase chain reaction.
‡Validation level for the reproducibility criterion: ⊕⊕ high intra- and inter-laboratory reproducibility confirmed; ⊕ high intra-laboratory reproducibility confirmed. Tests are validated as proposed by Meijer *et al* 2009[150].
$ validated on ThinPrep specimen, † evaluated on SurePath specimen.

11. HPV testing on self-samples

11.1 Retrieved studies

Searching the electronic databases, completed with hand searches, yielded 884 articles, from which 34 relevant papers were retrieved, evaluating the clinical accuracy of HPV DNA or RNA testing in self-samples[181,266,276-307]. Since one paper contained three separate studies[301], 36 studies were finally included in the meta-analysis, enrolling all together 154,556 women. In all but two selected studies[293,296], the comparator test was HPV testing on a clinician-sample. Non-reported accuracy data for the disease threshold CIN2+ were requested and obtained from three studies[297,301,61]. Contrary to the reviews included in the previous chapters, we included reports from diverse clinical settings to maximize the number of includable studies and to investigate the influence of the setting on the accuracy estimates. Sixteen studies concerned primary screening of generally healthy women [181,266,279,283,286,287,289,293,296,297,299,301,304,307], whereas in three others a high-risk population was screened[280,290,291]. In 17 reports, women referred to colposcopy were enrolled [276-278,281,282,284,285,288,292,294,295,298,300,302,303,305,306]. One study recruited women who were treated for cervical pre-cancer[294]. An overview of used HPV assays and sampling devices, which are grouped in five categories (brush, lavage, spatula, swab, and tampon) can be found in the Supplementary Tables of the published Lancet Oncology paper[308].

11.2. Assessment of study quality

Table 26 summarises the methodological quality of the 36 included studies, which was overall moderate to good with average negative scores for the 13 QUADAS items varying between 0% and 31%; equivocal scores varying between 0% and 58% and positive scores varying between 17% and 100%. Risk of bias regarding enrolment of subjects was considered low in 56%, moderate in 42% and high in 3% of studies. Reporting and execution of tests (description of cut-off, blinding of the index test towards comparator and reference test) was adequate in 72% and unclear in 36% but the risk of bias was never assessed as high. The quality of the verification with a reference standard (acceptable validity, blinding towards tests, avoidance of incorporating test results in final conclusion of disease outcome) was good in 89%, moderate in 8% and possibly problematic in 3% of the studies. The delay between self-sampling[9], clinician-sampling and verification with the reference standard was short in 69%, unreported in 25% and long in 6%. Partial verification was avoided in 78% but clearly present in 22%, whereas differential verification was absent in all but one studies. Withdrawal of patients was explained appropriately in 69% but not in 25% of the studies. In the majority of papers, non-interpretable results were poorly reported (56% and 61% for the evaluated tests and for the reference standard, respectively).

[9] Short delay was defined as ≤2 weeks between self- and clinician sampling or HPV testing on these samples and ≤6 months between the evaluated tests and the verification of the disease outcome.

11.3. Absolute sensitivity and specificity, by clinical setting

Table 27 presents the pooled absolute sensitivity and specificity of HPV DNA testing on self- and clinician-samples for detecting underlying CIN2+ and CIN3+ for each setting separately. Because of the substantial and significant variation by setting ($p<0.001$, data not shown) no overall pooled accuracy values were computed. The pooled sensitivity of HPV testing on self-samples collected in the frame work of primary screening was 76% (95% CI: 69-82%) and 84% (95% CI: 72-92%) for CIN2+ and CIN3+, respectively (see sROC plots in Figure 38). The pooled specificity for excluding CIN2+ was 86% (95% CI: 83-89%). HPV testing on clinician-samples showed high pooled sensitivity estimates for both CIN2+ (91%, 95% CI: 87-94%) and CIN3+ (95%, 95% CI: 91-97%), whereas the pooled specificity for <CIN2 was 88% (95% CI: 85-91%).

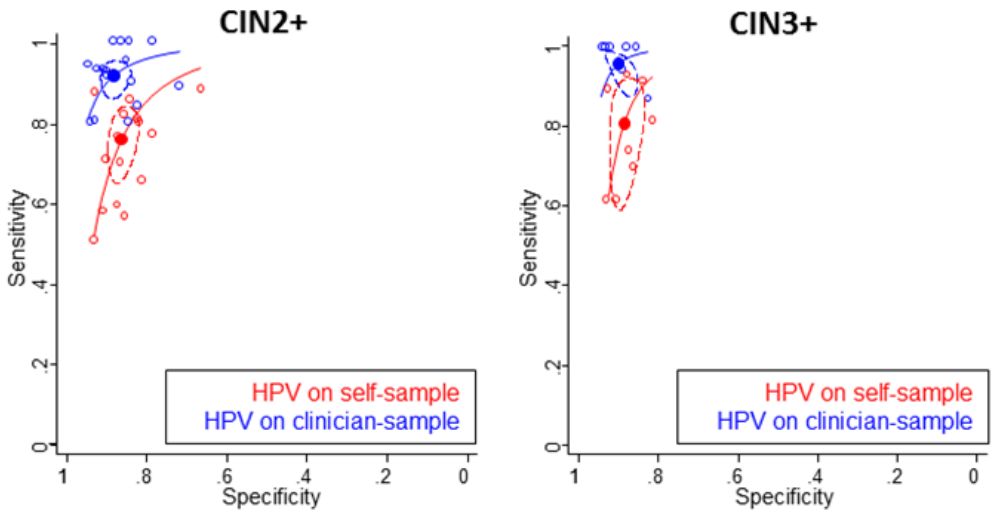


Figure 38. Meta-analysis of the accuracy in primary cervical cancer screening using HPV testing on self-samples (red) and HPV testing on clinician-samples (blue), for detection of CIN2+ (left) and CIN3+ (right). Hollow circles represent individual studies, the full curved line is the summary ROC curve, filled circles are the pooled accuracy measures, surrounded by the 95% confidence ellipse (interrupted line). Extracted from Arbyn et al, *Lancet Oncol* 2014[308].

Screening of high-risk women contributed only few observations. The accuracy estimates for CIN2+ of HPV testing in self- and clinicians-samples pooled from four observations were similar to those in primary screening (see Table 27). The accuracy for the outcome CIN3+ was more different from the pooled estimates in primary screening but the estimates in the high-risk testing group were based on only one study.

The pooled sensitivities for CIN2+ derived from follow-up studies were not substantially different from those observed in primary screening, at the exception of HPV testing of self-samples which showed a slightly higher pooled estimate (84%, 95% CI: 78-89%). On the contrary, the pooled specificity, in follow-up settings, usually was substantially lower with 56% (95% CI: 49-63%) and 58% (95% CI: 48-67%) for HPV testing in self- and clinician-samples, respectively (see Table 27).

11.4. Relative sensitivity and specificity of HPV testing in self-samples versus testing on clinician-samples

Table 28 shows the pooled relative accuracy measures of HPV testing on self- versus clinician-samples. From studies reporting results of multiple tests, the combination that yielded the highest relative sensitivity for CIN2+ was included in the meta-analysis under the condition that for a given study the HPV assay always had to be the same in self- and clinician-samples. The variation of the relative sensitivity and specificity of HPV testing on self- versus clinician-samples for detection of CIN2+ of all the included studies, grouped by clinical setting is shown in Figure 39. Substantial heterogeneity was identified among all studies (p always <0.001). However, there was no significant heterogeneity in relative sensitivity by clinical setting (p=0.345, data not shown). The variation in relative specificity by setting was also small (ratios of 0.98, 0.97 and 0.98, in primary screening, testing of high-risk groups and follow-up, respectively), but significant (p=0.001, data not shown) due to large numbers. Because of the rather limited inter-setting variation in the relative accuracy, meta-analytical pooling was performed across all clinical settings combined.
HPV testing on self-samples was, on average, significantly less sensitive than HPV-testing on clinician-samples, for both CIN2+ (ratio of 0.88 [95% CI: 0.85-0.91]; Figure 39) and for CIN3+ (ratio of 0.89 [95% CI: 0.83-0.96]). In addition, on self-samples, it was 4% less specific (ratio of 0.96, 95% CI:0.95-0.97) to exclude CIN2+ (Figure 39).

11.5. Effects of the HPV assay and of the self-sampling device

Distinguishing included studies by the general platform used for hrHPV testing revealed important differences. The relative sensitivity for CIN2+ on self- versus clinician-collected samples was significantly lower than unity when the HPV test was based on a principle of signal amplification. The ratios were 0.86 (95% CI 0.80-0.90) and 0.76 (0.70-0.83), for HC2 and Cervista, respectively. Also, the pooled specificity was significantly lower on self-samples when HC2 or Cervista were applied: pooled ratio of 0.96 (95% CI 0.94-0.98) (see Figure 40). However, when clinically validated PCR-based hrHPV assays, were used, no significant differences between self- and clinician samples was observed (see Figure 41). The APTIMA test was less sensitive (ratio = 0.64, 95% CI 0.46-0.90) but similarly specific (ratio = 0.99, 95% CI 0.98-1.01) on self- compared to clinician-taken samples.

No obvious collection device effects could be discerned in the relative sensitivity of HPV-testing on self- versus clinician-samples. With device effects, we mean that HPV testing with a validated PCR[10] on a self-sample taken with a given device yields a lower accuracy, in particular, lower sensitivity than HPV testing on sample taken by a clinician with a standard collection device. Since the included studies did not address comparisons between different self-sampling devices, and because of the variability in study design and settings, no strong conclusions about instruments for self-sampling can be drawn from the meta-analysis.

[10] A list of PCR-based HPV assay that is clinically validated for cervical cancer screening can be find in chapter 10.6.

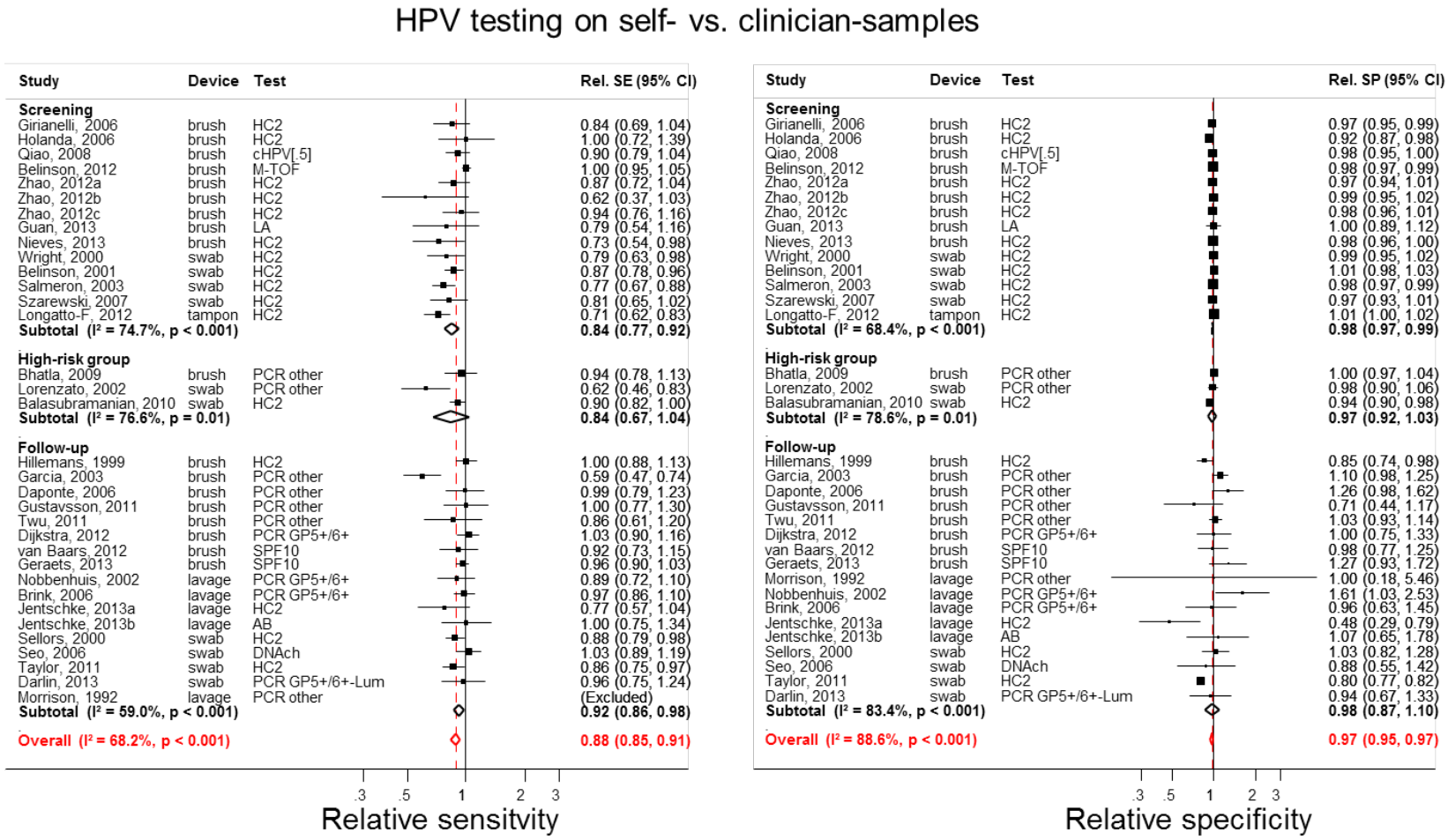


Figure 39. Relative sensitivity (left) and specificity (right) of HPV testing on self-samples compared to HPV testing on clinician-samples to detect CIN2+ or worse disease, by clinical setting. Derived from Arbyn et al, *Lancet Oncol* 2014[308].
p=test for inter-study heterogeneity, I^2= the percentage of total variation across studies due to inter-study heterogeneity. Rel. SE: relative sensitivity; Rel. SP: relative specificity.

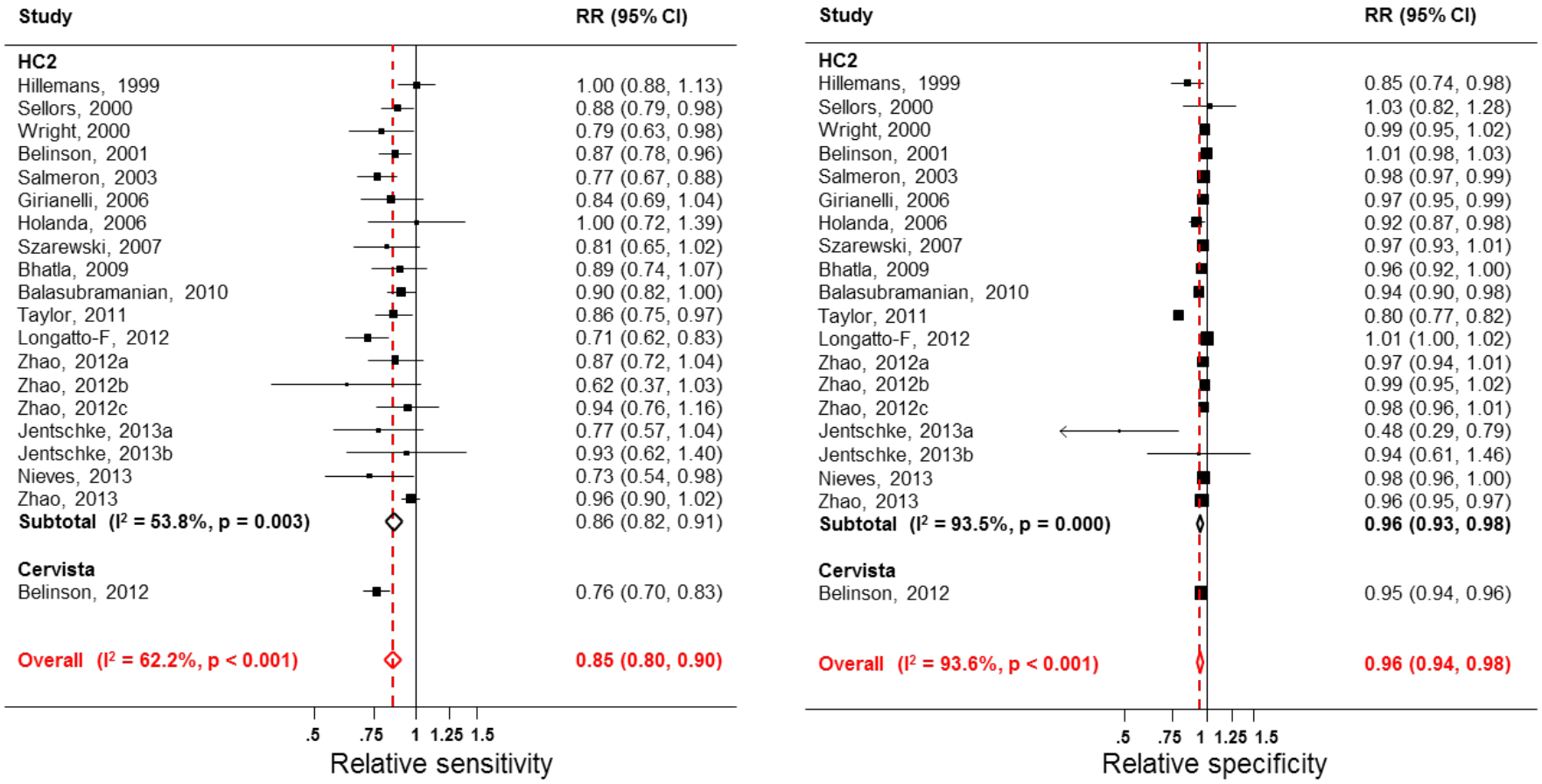


Figure 40. Relative sensitivity (left) and specificity (right) of high-risk HPV DNA testing, using validated signal amplification assays on self- versus clinician-collected samples to detect underlying cervical intraepithelial neoplasia of grade 2 or worse. Derived from Arbyn & Castle, CEBP 2015[309].

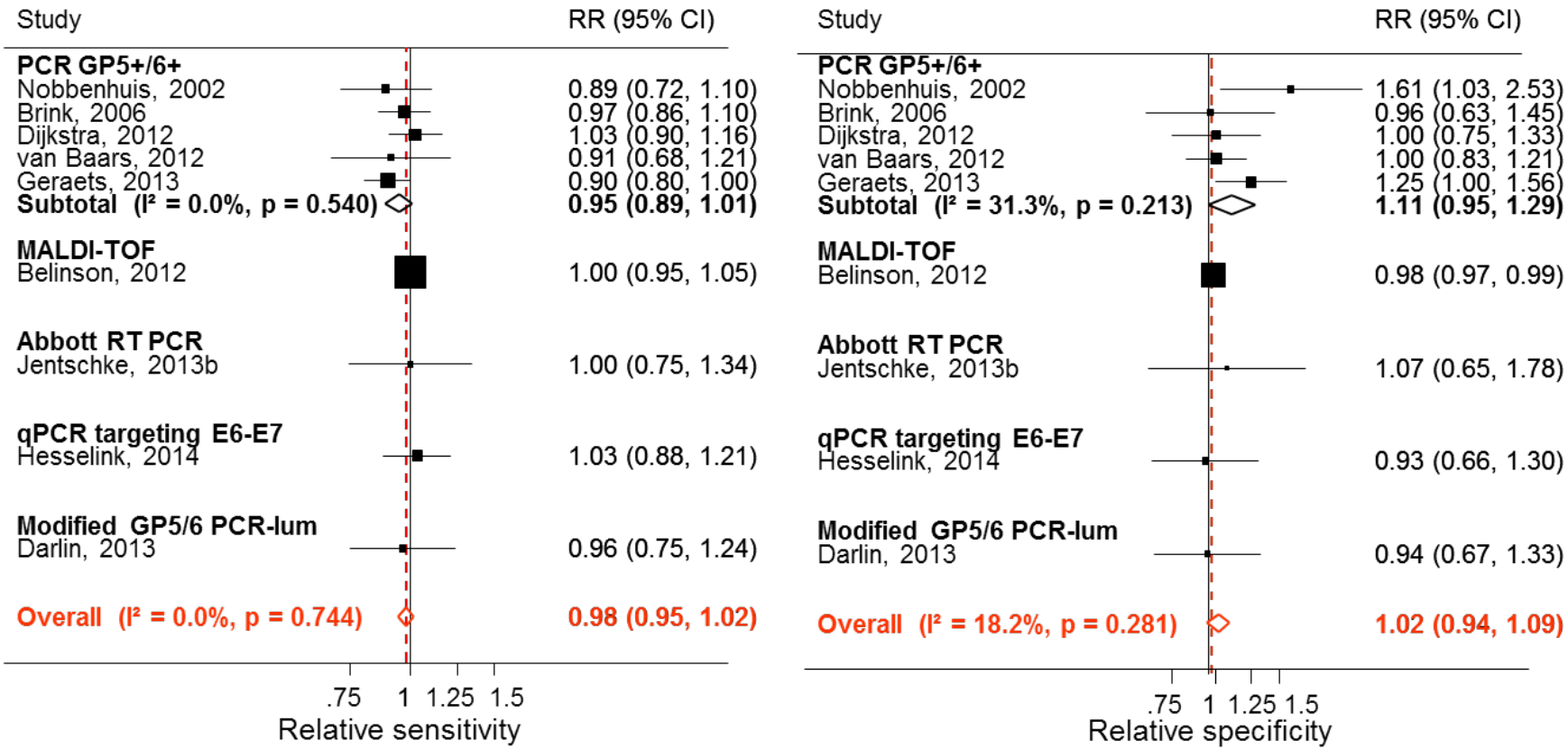


Figure 41. Relative sensitivity (left) and specificity (right) of high-risk HPV DNA testing, using validated PCRs on self- versus clinician-collected samples to detect underlying cervical intraepithelial neoplasia of grade 2 or worse. Derived from Arbyn & Castle, CEBP 2015[309].

11.6 Tables: HPV testing on self-samples

Table 26. Evaluation of the quality of each study according to the QUADAS-2 check list.

	Patient selection		Evaluated tests		Reference test			Flow & timing						Concerns of applicability		
Author, year	P1	P2	T1	T2	R1	R2	R3	F1	F2	F3	F4	F5	F6	Patient selection	Evaluated test	Reference test
Morrison, 1992	U	U	U	U	Y	U	Y	Y	Y	Y	Y	Y	Y	Moderate	Moderate	Low
Hillemanns, 1999	U	U	U	U	Y	U	Y	U	Y	Y	U	N	N	Moderate	Moderate	Low
Sellors, 2000	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Wright, 2000	U	U	Y	Y	Y	Y	Y	N	N	Y	N	N	N	Moderate	Low	Low
Belinson, 2001	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Lorenzatto, 2002	Y	Y	Y	U	Y	U	Y	Y	Y	Y	Y	Y	N	Low	Low	Low
Nobbenhuis, 2002	Y	U	U	U	Y	U	Y	Y	Y	Y	Y	Y	Y	Low	Moderate	Low
Garcia, 2003	Y	U	Y	U	Y	Y	Y	Y	Y	Y	N	Y	N	Low	Low	Low
Salmeron, 2003	U	U	Y	U	Y	N	Y	U	N	Y	Y	Y	Y	Moderate	Low	Moderate
Brink 2006	U	U	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	Moderate	Low	Low
Daponte, 2006	Y	U	U	U	Y	U	Y	U	Y	Y	U	N	N	Low	Moderate	Low
Girianelli, 2006	U	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Moderate	Low	Low
Holanda, 2006	U	U	Y	U	Y	Y	Y	Y	Y	Y	Y	N	N	Moderate	Low	Low
Seo, 2006	U	U	Y	U	Y	U	Y	Y	Y	Y	N	N	N	Moderate	Low	Low
Szarewski, 2007	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Qiao, 2008	Y°	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Bhatla, 2009	Y	U	Y	U	Y	U	Y	Y	Y	Y	Y	N	Y	Low	Low	Low
Balasubra., 2010	Y	U	Y	U	Y	U	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Gustavson, 2011	U	U	U	U	Y	Y	Y	Y	Y	Y	Y	N	N	Moderate	Moderate	Low
Lazc.-Ponce, 2011	Y	U	Y	U	Y	U	Y	U	N	Y	Y	N	N	Low	Low	Low
Taylor, 2011	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	Low	Low	Low
Twu, 2011	U	U	Y	U	Y	U	Y	U	N	Y	N	Y	Y	Moderate	Low	Low
Wikstrom, 2011	Y	U	Y	U	Y	U	Y	N	N	Y	Y	N	N	Low	Low	Low
Belinson, 2012	U	U	Y	U	Y	U	Y	Y	N	Y	Y	N	Y	Moderate	Low	Low
Dijkstra, 2012	U	U	U	Y	Y	U	Y	Y	Y	Y	N	N	N	Moderate	Low	Low
Long.-Filho, 2012	Y	U	Y	Y	Y	N	Y	U	Y	Y	Y	U	U	Low	Low	Moderate
van Baars, 2012	U	U	U	U	N†	U	U	Y	N	U	Y	N	N	Moderate	Moderate	High
Zhao, 2012a,b,c*	Y	Y	Y	U	Y	U	Y	Y	Y	Y	Y	N	N	Low	Low	Low
Darlin, 2013	U	U	U	U	Y	U	Y	Y	Y	Y	Y	Y	N	Moderate	Moderate	Low
Geraets, 2013	Y	U	U	U	Y	U	Y	Y	Y	Y	N	N	N	Low	Moderate	Low
Guan, 2013	N	U	U	U	Y	U	Y	Y	Y	Y	N	N	N	High	Moderate	Low
Jentschke, 2013a	U	U	U	U	Y	U	Y	U	Y	Y	Y	N	N	Moderate	Moderate	Low
Jentschke, 2013b	Y	U	Y	U	Y	U	Y	U	Y	Y	Y	N	N	Low	Low	Low
Nieves, 2013	Y	Y	U	Y	Y	N	Y	U	Y	Y	Y	Y	Y	Low	Low	Moderate

QUADAS items[129]: (P1) acceptable enrolment method, (P2) inappropriate exclusions avoided, (T1) pre-specified test cut-off, (T2) results of index and comparator tests blinded towards each other and reference test, (R1) acceptable reference test, (R2) results of reference test blinded towards index and comparator tests, (R3) incorporation bias avoided, (F1) acceptable delay between triage tests and reference test, (F2) partial verification avoided, (F3) differential verification avoided, (F4) withdrawals explained, (F5) non-interpretable results reported for tests, (F6) non-interpretable results reported for reference test. Each quality item is judged with: Y (fulfilled, green), U (unclear, yellow), N (not fulfilled, red).

°Participants recruited from randomly selected communes

†Only 44/134 women were verified histologically, the rest of the participants were verified cytologically

* Only data for SPOCCS III-1, III-2, and III-3 are included in this meta-analysis. Data also reported jointly by Belinson, *Int J Cancer* 2010 [310].

Table 27. Meta-analysis of the absolute sensitivity and specificity of HPV testing on self-samples and on clinician-samples to detect CIN2+ and CIN3+, by clinical setting (primary cervical cancer screening, testing of high-risk groups and follow-up of women because of previous cervical abnormalities).

Sample	**Outcome**	**No. studies** (test-device combinations)‡	**Sensitivity**	**Specificity**
Primary screening				
Self-sample	CN2+	14(16)	0.76 (0.69-0.82)	0.86 (0.83-0.89)
Clinician-sample		14(16)	0.91 (0.87-0.94)	0.88 (0.85-0.91)
Self-sample	CIN3+	6(8)	0.84 (0.72-0.92)	0.87 (0.84-0.90)
Clinician-sample		6(8)	0.95 (0.91-0.97)	0.89 (0.87-0.92)
Screening of high-risk group				
Self-sample	CN2+	3(4)	0.75 (0.58-0.87)	0.86 (0.77-0.92)
Clinician-sample		3(4)	0.88 (0.78-0.93)	0.88 (0.81-0.93)
Self-sample	CIN3+	1	0.42 (0.27-0.57)†	0.81 (0.76-0.87)†
Clinician-sample		1	0.80 (0.67-0.93)†	0.82 (0.77-0.88)†
Follow-up				
Self-sample	CN2+	17(19)	0.84 (0.78-0.89)	0.56 (0.49-0.63)
Clinician-sample		17(19)	0.91 (0.86-0.94)	0.58 (0.48-0.67)
Self-sample	CIN3+	5(7)	0.85 (0.76-0.91)	0.45 (0.36-0.54)
Clinician-sample		5(7)	0.96 (0.92-1.00)*	0.46 (0.35-0.57)*

‡In some studies, more than one HPV assay was used, in this situation, the number of test-collection device combinations, entered in the meta-analysis, is mentioned between brackets; * sensitivity and specificity pooled separately because lack of fit of the bivariate normal model;
† no meta-analytical pooling since only one study;
CIN2/3+: cervical intraepithelial neoplasia of grade II/III or worse; HPV: assay that picks up high-risk types of the human papillomavirus.

Table 28. Pooled relative sensitivity and specificity of HPV testing on a self- versus on a clinician-sample, to detect CIN2+ and CIN3+.

Outcome	**No. studies**	**Sensitivity ratio (95% CI)**	**Specificity ratio (95% CI)**
CIN2+	34	0.88 (0.85-0.91)*	0.96 (0.95-0.97)*
CIN3+	12	0.89 (0.83-0.96)*	0.96 (0.93-0.99)*

* statistically significant.
CIN2/3+: cervical intraepithelial neoplasia of grade II/III or worse

12. Discussion

12.1. Triage of women with ASC-US

The meta-analysis on the triage of ASC-US presented here, demonstrated a significantly improved sensitivity of HC2 compared to repeat cytology at cut-off ASC-US (27% and 14% increase for CIN2+ and CIN3+, respectively), which was coupled with a non-inferior specificity. These findings corroborate conclusions formulated in our previous reviews[93,94,311]. As a result, hrHPV testing for triage of ASC-US cytology was recommended and widely applied in the United States and Europe[223,312].

Three other HPV DNA-based assays (Abbott PCR, Papillocheck, and Cervista) and one mRNA-based assay (APTIMA) were found also to be appropriate triage tests for women with a diagnosis of ASC-US. By extension all other hrHPV DNA tests that fulfil requirements for primary cervical cancer screening can be accepted as well for triage of ASC-US (see sub-chapter 10.6).

Nonetheless, the absolute specificity of hrHPV DNA-based assays for exclusion of CIN2+ still is suboptimal (range42-60%) indicating that there is room for developing more specific triage assays . Remarkable is the higher specificity of RNA- compared to DNA-based HPV assays. The higher specificity of the Pretect Proofer assay can be explained easily since this test only targets RNA of five HPV types. Consequently, the sensitivity of Pretect Proofer was significantly lower than DNA tests targeting all hrHPV types. The explanation, why APTIMA (that detects RNA of 14 hrHPV types) is more specific but as sensitive compared to most hrHPV DNA tests, is more controversial. The higher specificity might be due to the fact that presence of the transcripts of the *E6* and *E7* oncogenes indicates that the virus is triggering neoplastic transformation, whereas presence of hrHPV DNA just indicates that the virus is there. Certain experts have remarked that RNA is not necessarily a marker of carcinogenesis. It could be that APTIMA just is analytically less sensitive than current hrHPV DNA tests and picks up less transient infections. hrHPV DNA assays could be made as specific as APTIMA by a higher viral load cut-off. Whether the woman with a negative APTIMA test has a similarly low risk to develop in the next five years a CIN3 lesion or a cancer as a woman with a negative hrHPV DNA test is further discussed in § 12.6.2.

Single $p16^{INK4a}$ and double p16/Ki-67 immuno-staining were shown to meet the criteria for good specificity without loss of significant loss in sensitivity, compared to HC2. Indeed, in a previous systematic review, p16 immuno-staining had demonstrated promising results in terms of lowering the number of false referrals, compared to HC2[313,314]. In this report and a previous meta-analysis[121], based on data from 8 studies, the specificity of $p16^{INK4a}$ to reject CIN2+ was 1.80 times higher than that of HC2, without losing sensitivity. In addition, p16/Ki-67 double immuno-staining demonstrated a specificity which was twice as high as for HC2. However, these findings were based on only one study[315]. The meta-analysis on the accuracy of double staining for p16/Ki-67, including studies up to August 2017, is currently being updated at the Unit Cancer Epidemiology.

12.2. Triage of women with LSIL

In the triage of women with LSIL, HC2 demonstrated a significantly improved potential to detect CIN2+ (relative sensitivity: 1.23) compared to cytology. For CIN3+, however, the sensitivity of HC2 was not significantly better than that of repeat cytology (cut-off ASC-US). Moreover, a substantial loss in specificity (-34% for CIN2+, -44% for CIN3+) was observed, potentially leading to over-diagnosis and over-treatment.
It is generally believed that LSIL in fact is the manifestation of a productive HPV infection with low potential of neoplastic transformation[68]. In the American ASCUS-LSIL Triage Study (ALTS), 75-80% of women with LSIL were confirmed to harbour hrHPV types[316]. In the same way, our meta-analysis demonstrated that 79.5% (95% CI=77.1-81.7%) of women with LSIL had a positive HC2 test, while underlying CIN2+ was present in only 21%. Given its high positivity rate, HC2 has limited capacity to distinguish between cases with and without clinically significant disease.
Due to poor values for specificity, yielding a PPV which is hardly different from the pretest probability of high-grade CIN, the American Society for Colposcopy and Cervical Pathology (ASCCP) does not recommend reflex HPV triage, but proposes to refer to colposcopy. If colposcopy and/or biopsy are normal or only reveal CIN1, an HPV test at 12 months or two repeat smears after the initial LSIL smear are recommended. In The Netherlands, Bais and Berkhof showed that delayed HPV and repeat cytology testing in patients with mild dyskaryosis after 6 and 18 months is both safe and more cost-effective than immediate HPV triage [317,318]. Postponing triage, allows viral clearance, which over a period of 6–12 months can vary from 18% to 45%[317] and therefore reduces the need for colposcopy. However, this strategy requires good compliance with follow-up recommendations.

hrHPV tests, other than HC2, generally demonstrate a good sensitivity with approximately half of them meeting the requirement of a lower 95% confidence bound of at least 0.90 (relative sensitivity compared to HC2). Of the assays that met the sensitivity criterion, Abbott RT-PCR and APTIMA satisfied the requirements for specificity (relative specificity $\geq$1.10 and lower 95% confidence bound $\geq$ 1.00) as well. Testing with biomarkers or identifying mRNA of the 5 main HPV types or DNA of HPV16/18 showed substantially improved specificity but did not reach the minimal sensitivity.

Over-expression of p16 or p16/ki67, identified by immunocytochemistry is an interesting marker. A recent network meta-analysis including hrHPV DNA, hrHPV RNA and p16 assays showed that the latter was the most efficient (best compromise between sensitivity and specificity) of all triage markers. However, as shown in the next sub-chapter (Figure 43), a negative p16/Ki67 test does not allow to release the patient and still requires further follow-up.

12.3. Utility of HPV testing in a given setting derived from PPP plots

The concept of pretest-posttest probability (PPP) plots as a tool to display the utility of a test, was presented recently, in a meta-analysis of the accuracy of HPV16/18 genotyping to triage women with minor cervical abnormalities[145]. Meta-analyses of DTA studies generate the average sensitivity and specificity as main result together with a framework to assess the quality of included studies that may explain the observed inter-study heterogeneity[137]. However, the probability that a patient has the suspected disease for those with a positive or negative test, is not immediately derivable from the pooled accuracy estimates since this post-test probability depends on the clinical context. The purpose of a triage test is to stratify risk and to facilitate decision making according to benchmarks defined by experts. An optimal test distinguishes those who need further diagnostic or therapeutic work-up when the test is positive from those with a negative triage test who can be released to the next screening round. The PPP plot connects the pretest risk of disease on the left Y-axis with the posttest risks on the right Y-axis. The PPP plots contain colour bends suggesting appropriated management, based on decision thresholds determined by experts. Green defines the zone where a patient can be released to the routine screening programme, the red zone implicates a need for further management, whereas the in-between yellow suggests further surveillance. The steeper the slope of the connecting lines linking the pretest to the posttest probability of pre-cancer, the more efficient the triage. The PPP plot displays the change in risk according to the result of one test or a of combination of tests. By using log(Odds Probability) scale on the vertical axes, the angles between pre- and posttest risks directly reflect the logarithm of the likelihood ratios. Several examples of triage situations are described in Table 29 and Figure 42-Figure 44.

Table 29. Accuracy (sensitivity, specificity, likelihood ratios), pre-test and post-test probabilities of CIN3+ for triage tests applied on women with ASC-US, LSIL, ASC-H, hrHPV+ ASC-US or LSI, or hrHPV+ identified through screening. Used tests are HC2, genotyping for HPV16/18, p16 with or without Ki67, cytology at cut-off ASC-US or worse.

		Pooled					Posttest probability	
Application	**Test**	**Sensitivity**	**specificity**	**Pretest probability**	**PLR**	**NLR**	**if test+ (PPV)**	**if test- (cNPV)**
Triage ASC-US	HC2	94%	50%	7%†	1.90	0.11	12%	0.8%
Triage LSIL		98%	25%	9%†	1.31	0.08	11%	0.8%
Triage ASC-H		91%	42%	20%‡	1.57	0.21	28%	5.1%
Triage ASC-US	p16/Ki67	84%	60%	7%†	2.10	0.27	13%	1.9%
Triage LSIL		92%	47%	9%†	1.74	0.17	14%	1.6%
Triage ASC-US	HPV16/18†	74%	62%	6%*	3.23	0.38	17%	2.3%
Triage LSIL		73%	61%	9%*	2.55	0.41	19%	3.7%
Triage hrHPV+ ASC-US		82%	72%	10%*	1.92	0.43	18%	4.6%
Triage hrHPV+ LSIL		96%	57%	11%*	1.85	0.45	18%	5.1%
Triage of HPV+	Cyto, 1x	90%	53%	9%**	2.93	0.25	22%	2.4%
	Cyto, 2x	85%	65%	9%**	2.23	0.07	18%	0.7%
	Cyto & HPV16/18	94%	50%	9%**	1.90	0.20	16%	1.9%
	p16 or p16/Ki67	98%	25%	9%**	2.43	0.23	19%	2.2%

PPV: positive predictive value = risk of CIN3+ if the triage test is positive; cNPV: complement of the negative predictive value = 1-NPV = risk of CIN3+ if the triage test is negative; PLR: positive likelihood ratio; NLR: negative likelihood ratio.

Source:† Arbyn et al, *Vaccine* 2012[105]; ‡ Xu et al, *Cancer Cytopathol* 2015[319]; * Arbyn et al, A*nn Intern Med* 2015[145]; ** Arbyn et al, *KCE* 2015[111].

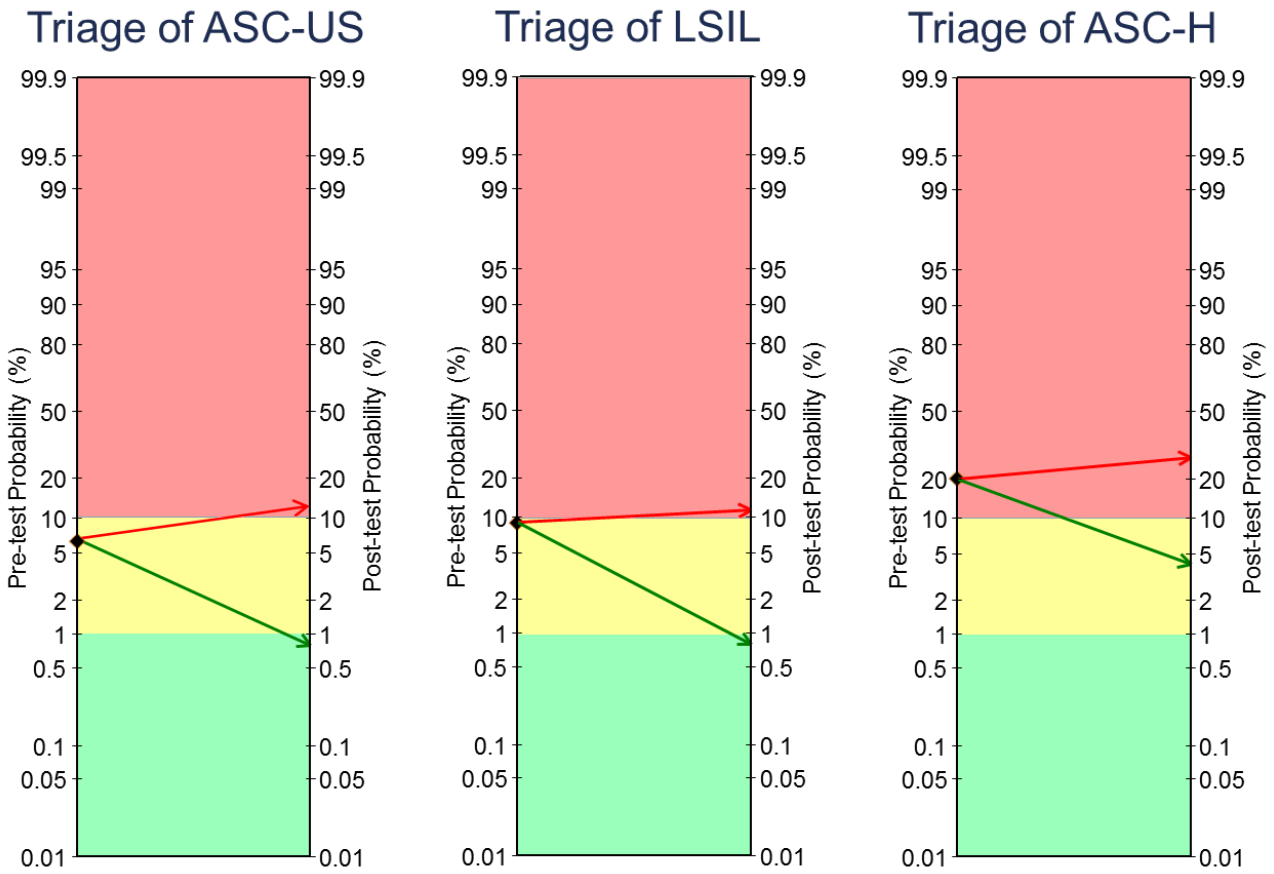


Figure 42. PPP plot connecting the pretest and posttest probability for CIN3+ when the HC2 assay is used to triage women with ASC-US (left), LSIL (middle) or ASC-H (right).
ASC-US: atypical squamous cells of undetermined significance; ASC-H: atypical squamous cells where high-grade lesions cannot be excluded, LSIL: low-grade squamous intraepithelial cells, PPP: pretest posttest probability.

The probability of disease when the test is positive corresponds with the positive predictive value, whereas the probability of disease when the test is negative corresponds with the complement of the negative predictive value (cNPV = 1 – NPV).
In case of ASC-US triage, a shift is observed from the yellow-zone (pre-test) to the red and green zones if the test is positive and negative, respectively, indicating a useful triage application.
When women with LSIL are triaged with HC2, the post-risk for HC2+ women (PPV=11%) is hardly higher than the pretest CIN3+ risk (9%), indicating a poorly efficient positive triage process. However, LSIL women with a negative HC2 result have a similarly low risk as hrHPV negative women with ASC-US (cNPV<1%).
In women with ASC-H, a positive HC2 test does not change the decision to refer to colposcopy, since the pretest CIN3+ risk exceeds already 10%. A negative HC2 does not allow releasing women with ASC-H but could generate a recommendation to keep them under surveillance, since the posttest probability of CIN3+ screening (cNPV=5%) is in the yellow zone.
Triage with a more specific biomarker set, such as p16/ Ki67 cyto-immunochemistry, allows more efficient triage, not only in case of ASC-US but also for LSIL, (PPV for CIN3+ of 13% and 14%, respectively, see Figure 43). Nevertheless, for both ASC-US and LSIL, the posttest CIN3+ risk associated with a double stain negative triage result, exceeds 1%, suggesting a recommendation of further surveillance.

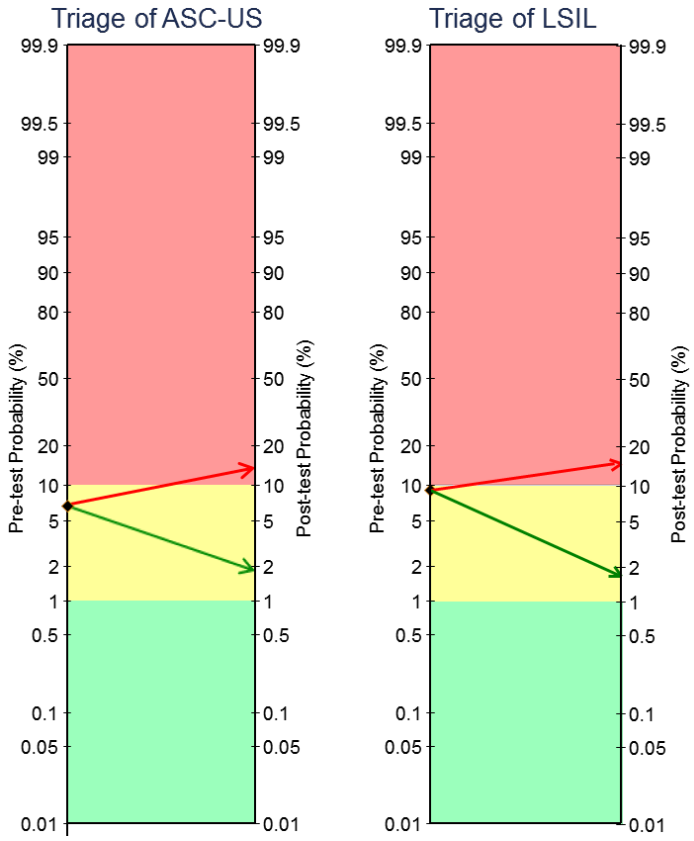


Figure 43. PPP plot connecting the pretest and posttest probability for CIN3+ when the p16/Ki67 double staining is used to triage women with ASC-US (left) or LSIL (right).
ASC-US: atypical squamous cells of undetermined significance;, LSIL: low-grade squamous intraepithelial cells, PPP: pretest posttest probability.

Triage of women with ASC-US or LSIL with HPV16/18 genotyping is insufficiently sensitive (~70%) to be used as a single triage test (Table 29). However, HPV16/18 genotyping could be used as a second test to triage those who are hrHPV+ at the first step triage (see Figure 44).

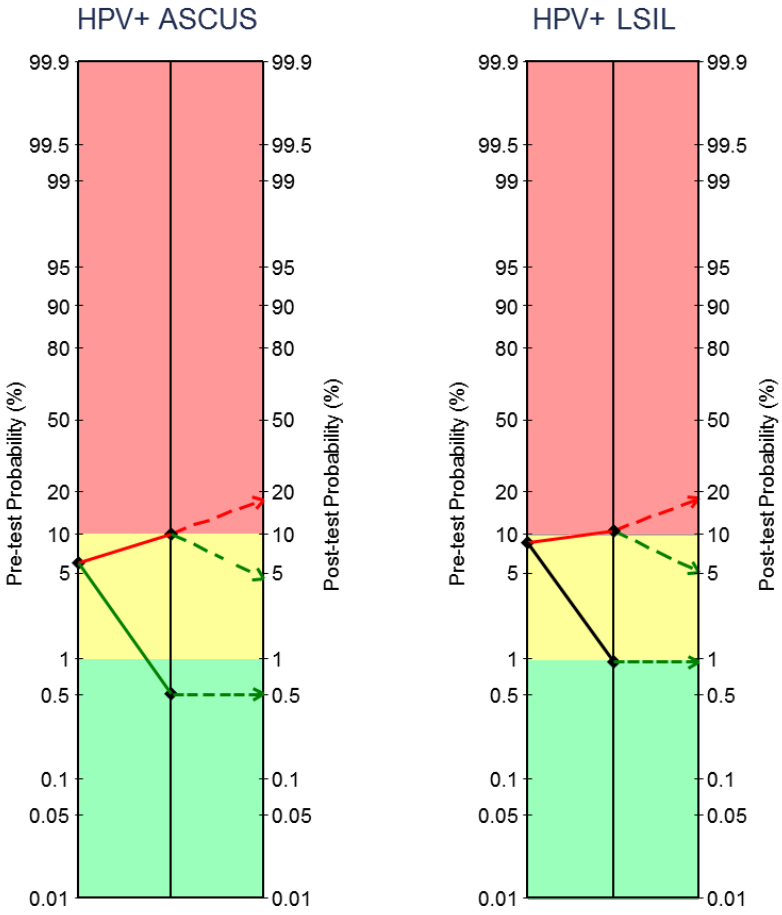


Figure 44. PPP plot connecting the pretest and posttest probability for CIN3+ when HPV16/18 genotyping is used to triage women with ASC-US (left) or LSIL (right), who are hrHPV+ at a first triage step.

For a discussion of strategies usable to triage women who test HPV+ in primary screening (bottom of Table 29), see sub-chapter 10.3.

12.4. Accuracy of HPV-DNA testing and cytology to predict treatment failure

.

Women treated for CIN2/3 by excisional or ablative procedures still have an increased risk of residual or recurrent CIN2 or worse disease compared to the general population[103,173,320]. Therefore, further surveillance is crucial. Although a rapid decline in the occurrence of residual or recurrent CIN2 and CIN3 is observed after the first two years post-treatment (see Figure 1), long-term follow-up data demonstrate that the risk for invasive cancer remains elevated, even up to 20 years[99,173].

The rate of residual or recurrent high-grade CIN, observed within a period of 18 months or more post-treatment, varied among the included studies from 4% to 16%, with a pooled average of 8%. This residual/recurrent disease was predicted by hrHPV-DNA testing with a considerably higher sensitivity (ratio 1.29, 95% CI 1.18-1.40), but a slightly lower specificity (ratio 0.94, 95% CI 0.90-0.99), compared to cytology. Combined testing for hrHPV-DNA and cytology, was slightly but not significantly more sensitive (ratio 1.07, 95% CI: 0.97-1.17) but resulted in a significant drop of the specificity (ratio 0.93, 95% CI0.88-0.98), compared to hrHPV-DNA testing alone.

The significant heterogeneity which was observed between the included studies, may be explained by diverse study characteristics. Studies differed by duration of follow-up. Furthermore, variations in time of testing after treatment were observed. For the analysis presented here, only data of tests that had been performed within 3 to 9 months after treatment were included. Additionally, the age of the participants and other patient-related characteristics might account for the observed variation. Nonetheless, the data presented here, demonstrate a consistent and largely significant increase in sensitivity when using hrHPV-DNA testing compared to cytology follow-up. Furthermore, although inclusion criteria are not always identical, our conclusions confirm those from previous systematic reviews regarding sensitivity. However, the specificity of HPV was significantly lower than that of cytology, in our meta-analysis, whereas previous reviews, including less studies, found a relative specificity that did not differ from unity[68,69,119].

Co-testing provided an insignificant gain in sensitivity (7%) compared to hrHPV-DNA testing alone but resulted also in a loss of specificity (-7%, 95% CI -2 to -12%). Given the setting (post-treatment follow-up) where higher priority may be given to sensitivity than specificity, clinicians may prefer co-testing rather than HPV testing alone.

The number and timing of follow-up tests are important issues to be defined. One single testing, even with a hrHPV-DNA assay, might not be sufficient to identify women at risk for residual or recurrent high-grade disease[321,322]. A study that assessed the predictive value of consecutive testing over 10 years after treatment, demonstrated that one negative test at 6 months after treatment resulted in a 10-year risk of CIN3+ of 2.1% (if hrHPV-negative), 2.8% (if cytology-negative) and 1.4% (if combined hrHPV- and cytology-negative)[322]. By adding one or two follow-up tests however, safety could be improved considerably. For example, three consecutive negative cytology results (6, 12 and 24 months) or two negative co-tests (hrHPV-negative and cytology-negative, at 6 months and 12 months) were associated with a low risk of residual CIN3+ of 0.7% and 0.0%, respectively, which was not different from the general population with negative cytology

at screening. On the other hand, one positive test at 6 months increased the risk of CIN3+ dramatically: 29%, 13% and 23%, if hrHPV-positive, cytology ASC-US+, or one of both tests positive, respectively[322]. Based on this longitudinal evidence, authors recommended co-testing (cytology and hrHPV) at 6 and 18 months after treatment with referral to colposcopy if one these tests is positive and referral to the normal screening schedule if women tested negative for all tests[322].

12.5. Accuracy of HPV testing in cervical cancer screening

The updated meta-analyses corroborates previous reviews concluding that hrHPV testing is substantially more sensitive than cytology in identifying underlying CIN2+ and CIN3+ in the context of primary cervical cancer screening[69]. However, one drawback is the lower specificity. The wide spread of the accuracy estimates in certain developing countries may be explained by variability in the quality of the applied reference standard. Further variation is observed according to the intensity of verification and methods to adjust for verification bias [104,323].

It should be noted that the higher cross-sectional sensitivity of hrHPV testing for detecting CIN2+ and CIN3+ provides insufficient evidence that HPV-based screening will decrease the incidence of cervical cancer more than cytology-based screening. Most CIN2 and CIN3 lesions clear and it cannot be excluded from cross-sectional studies that HPV tests just pick up more regressive disease[223]. For, this reason, hrHPV-based cervical cancer screening was not recommended yet in the 2nd edition of the European Guidelines for Cervical Cancer Screening, considering evidence available in 2006[223].
Besides cytology, genotyping for HPV type 16 or 16 and 18, combinations of HPV16/18 genotyping and cytology, or HPV retesting to detect persistent infections can stratify the risk for CIN3+ and can be used in triage of hrHPV-positive women [204,205,324,325]. The 10-year cumulative risk associated with HPV16, HPV18 or other high-risk HPV infection, among women included in the Portland study was 17%, 14% and 3%, respectively [326]. In the Athena, trial, identification of HPV16 or HPV18 with the Cobas-4800 test showed similar sensitivity and PPV for CIN3+ (60% and 16%, respectively) as cytology at cut-off ASC-US+ (53% and 14%, respectively) to triage hrHPV positive women [208]. Other candidate markers which could be considered for triage are: testing for viral RNA of all or a limited number of hrHPV types[327], p16INK4a immunocytochemistry or p16INK4a/Ki67 double staining[121]. However, until present, evidence on the accuracy of these markers in triage of hrHPV positive women is still insufficient.
Currently, no marker or combination of markers allow perfect reflex triage. hrHPV positive women being cytology negative or HPV16/18 negative must be kept under surveillance and invited for repeat testing. Also primary HPV screening using a more specific test such as the APTIMA RNA assay[328] or Hybrid Capture-2 at a higher viral load cut-off[329] increases PPV with no or a small losses in cross-sectional sensitivity. However, longitudinal data are needed before such policies can be recommended (see sub-chapter 12.6.2).

As demonstrated in sub-chapter 12.3 (see Table 29), PPP plots can also be used to appreciate the utility of algorithms considered to triage HPV+ women (see Figure 45).
The second PPP plot of Figure 45 confirms that two subsequent cytology tests are associated with CIN3+ lower than 1%. This level of safety could not be reached with one reflex test or test combination.

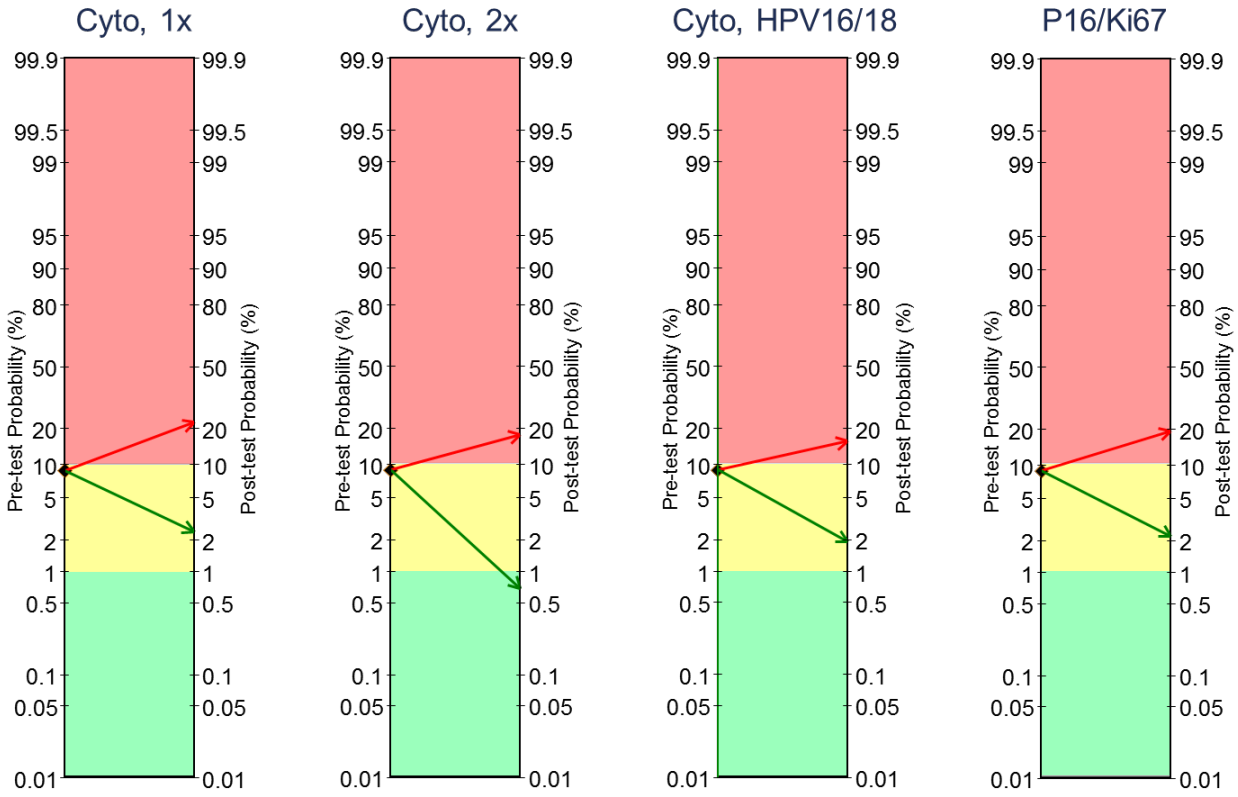


Figure 45. PPP plots connecting the pretest and posttest CIN3+ risk for four strategies to triage hrHPV+ women: 1) reflex cytology at cut-off ASC-US+; 2) reflex cytology at cut-off ASC-US+ followed by a second cytology examination if reflex cytology was negative; 3) reflex cytology at cut-off ASC-US+ combined with genotyping for HPV16/18; 4) double staining for p16 and Ki67.

12.6. Effectiveness of HPV testing in cervical cancer screening

12.6.1. Existence of HPV-negative cervical cancers: a reason to prefer co-testing?

A recent report from a large cytopathology laboratory in the USA, linking prior testing of cervical cell specimen with the subsequent occurrence of invasive cervical cancer revealed significantly higher false-negativity rates for HPV testing (19%) than for the combination of HPV testing and cytology (6%), which corresponds to a sensitivity for the outcome cancer of 81% and 94%, respectively[330]. The authors concluded that co-testing should be the preferred screening policy. To verify whether the sensitivity of the recent reports published in Cytopathology, 2015[330-332] were exceptional or not, we extended our

meta-analyses on the accuracy of hrHPV tests for CIN2+ or CIN3+ in cervical screening, summarized in sub-chapter 10.1 and verified the sensitivity for the outcome cancer and included also more recent reports. The sensitivity of co-testing and HPV-only testing for cancer pooled from 17 studies was 97% (95% CI 94-99%) and 87% (95% CI 83-91%), respectively (Figure 46). The absolute sensitivity varied significantly among studies ($p<0.001$, $I^2>65\%$), and depended on verification procedures and the interval between testing and cancer diagnosis. However the relative sensitivity did not show significant inter-study heterogeneity (p=0.13) (Figure 47).

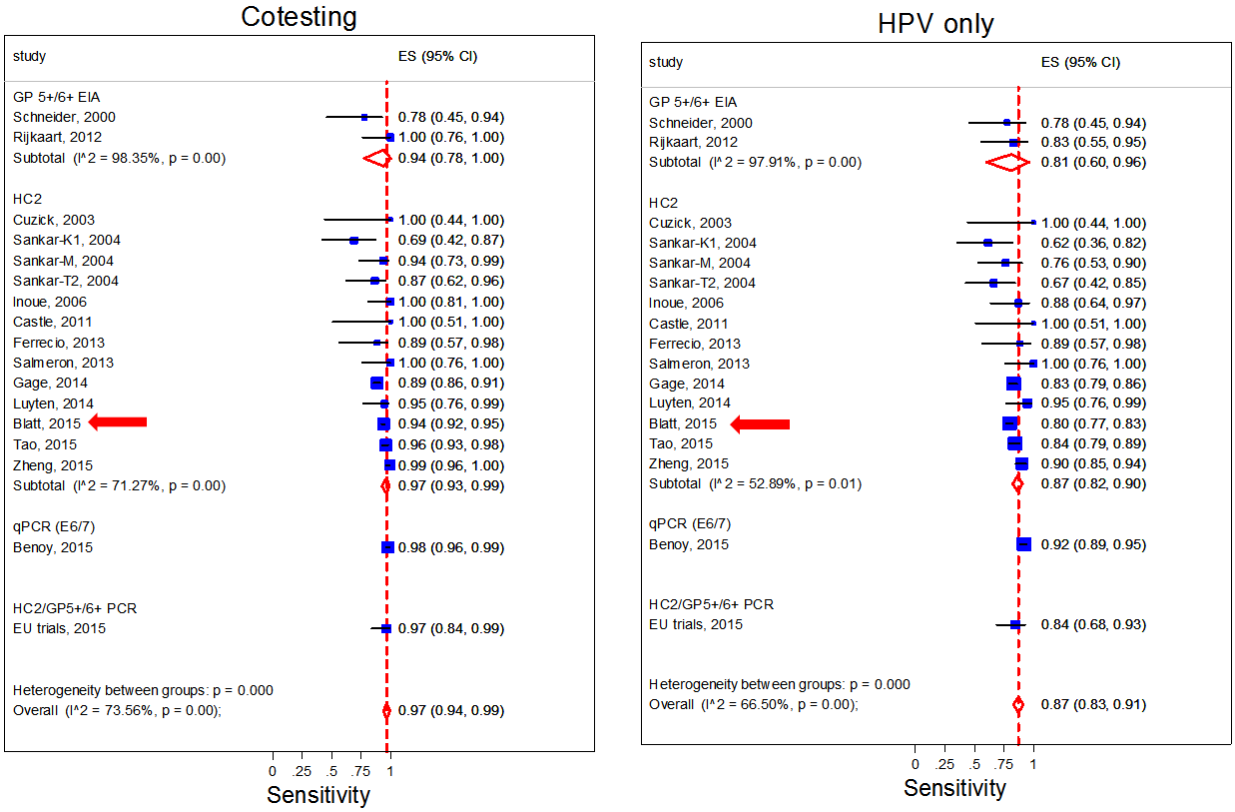


Figure 46. Meta-analysis of the sensitivity for the outcome cancer in cervical cancer screening with co-testing (HPV assays combined with cytology, at left) or with HPV assays only (at right). The red arrow points to the findings of Blatt et al, *Cancer Cytopathol* 2015[330].

On average co-testing was 1.10 (95% CI 1.07-1.13) times more positive in cancers than HPV testing, which was lower than the relative sensitivity observed by Blatt et al (1.18, 95% CI 1.13-1.23). Although omission of the latter study from the meta-analysis decreased the pooled relative sensitivity to 1.08 (95% CI 1.06-1.11), co-testing still remained significantly more sensitive than HPV-only-testing.

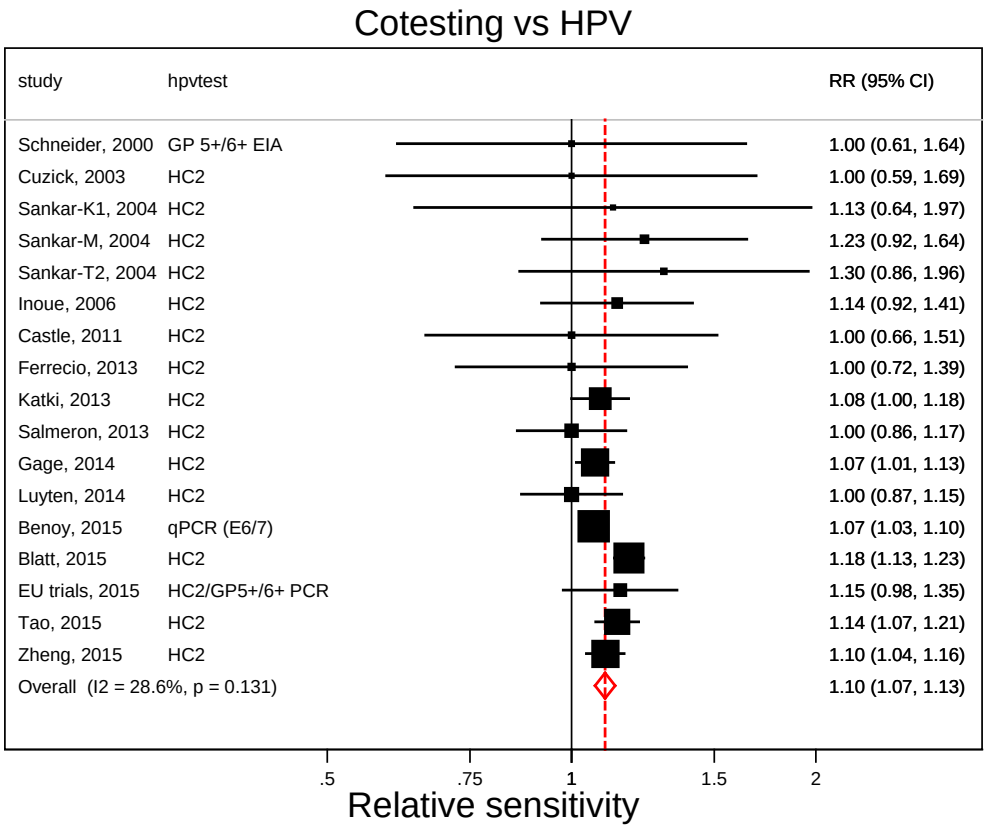


Figure 47. Meta-analysis of the relative cross-sectional sensitivity for cancer of co-testing vs. HPV-only-testing in cervical cancer screening.

12.6.2. Longitudinal criteria to demonstrate effectiveness of other screening tests which are not HPV DNA tests

The longitudinal evidence, presented in subchapter 10.2, was pivotal for the development of guidelines recommending cervical cancer screening with HPV testing at longer intervals than with cytology-based screening[105,199]. Most often, evidence regarding accuracy of tests or efficacy or effectiveness of interventions is reduced to dichotomized outcome measures. Pooling of dichotomized effectiveness measures derived from individual studies irrespective of the timing that events occur, may be problematic when outcomes are assessed over different follow-up periods. Restriction of meta-analyses to studies with the same follow-up or separating them by period follow-up, compromises the synthesizing power, unless individual-person-data files can be obtained.

Currently a new statistical project, called "*metaSURV*" is currently initiated at the Unit of Cancer Epidemiology (WIV-ISP, Brussels), aiming to pool time-to-event data including digitized Kaplan-Meier curves, extracted from published studies, as a new promising exploratory method. Multiple survival analysis models are being explored which account for availability or lack of accompanying data and fitting conditions[333].

Figure 38 shows the Kaplan Meier curves representing the cumulative incidence of CIN3+ after a negative conventional cytology result, after a negative HPV test and after a negative co-test, obtained by digitization of the figures from two published papers[211,334].

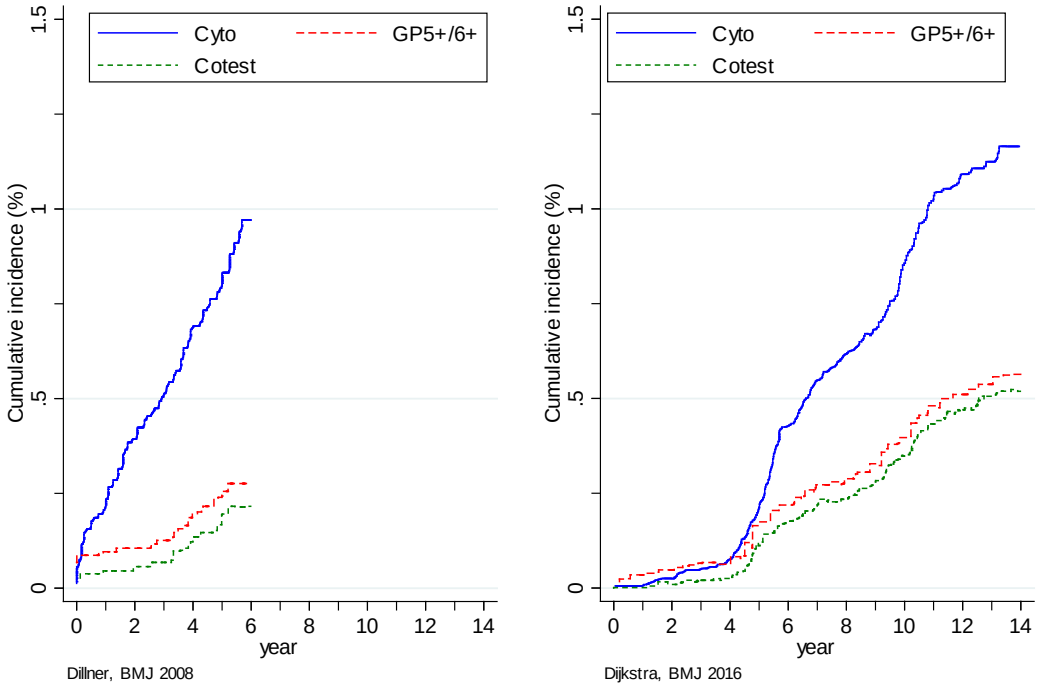


Figure 48. Cumulative incidence of CIN3+ after a negative screening test result: a) after negative conventional cytology (blue), b) after a negative hrHPV DNA test result (GP5+/6+PCR or HC2) (red) and c) after a negative co-test, as observed in a pooled series of European screening studies (left, Dillner, BMJ 2008[211]) and in the Dutch randomised trial (right, Dijkstra, BMJ 2016[334]). Logrank test on simulated data behind cumulative incidence curves: p for differences between cytology (blue) and HPV (red): 0.0004 (Dillner) and <0.0001 (Dijkstra); p for differences between HPV (red) and co-testing (green): 0.40 (Dillner) and 0.74 (Dijkstra).

The cumulative incidence after a negative cytology screen test result (blue curves in Figure 48) increased more rapidly over time than those after a negative HPV (red curve) or co-test (green curve) result. The blue curve in the multi-centric European study (Dillner, 2008) increased more quickly than the one in the Dutch trial (Dijkstra, 2016), since in former study, where retesting after negative screening was done more frequently in the framework of opportunistic screening. In both studies, the green curves (after negative co-testing) were situated slightly lower than the red curves (after negative-HPV only testing).
The vertical distances between differently coloured curves reflect the differences in longitudinal sensitivity between the baseline tests. The relative risk or ratio of the cumulative incidence after a negative test vs. the cumulative incidence after a negative result of a more sensitive test indicates the magnitude of the higher risk of disease missed by the former test. In Figure 49, these relative risks of CIN3+ and of cancer after a negative HPV test versus after a negative co-test is plotted for studies documenting outcomes after three, five, ten and fourteen years of follow-up. After 3 and 5 years, significantly more CIN3+ and cancer cases were detected after a negative HPV test than after a negative co-test (with RRs being 1.25 and 1.35 for outcome CIN3+, and 1.69 and 1.31 for outcome cancer). These data indicate, with more convincing longitudinal evidence that co-testing on average protects more effectively against future CIN3+ and even against cancer, than the cross-sectional accuracy data presented in sub-chapter10.1. However, in terms of differences in absolute risks (also shown in Figure 49), the differences are small and must be balanced against the huge additional costs associated with co-testing, compromising the cost-effectiveness of co-testing policies. Remarkable is the additive effect: small and rather constant distance between the green and red curves. This may be explained by the rather small number of prevalent CIN3 and cancer HPV that were HPV negative but cytology positive. There is no evidence that these cases were prevented by co-testing. Only randomised trials comparing HPV alone screening with screening by cotesting could provide the evidence on improved efficacy of the latter.

The cumulative incidences (including differences and ratios), as displayed graphically in Figure 48 and Figure 49 may inform decision makers about the choice of a screening policy.

The relative cumulative incidences could only be computed from published tables, for a limited number of follow-up periods. The advantage of digitizing Kaplan-Meier curves is that more data can be generated from more studies and more time periods can be considered, irrespective of the actual follow-up duration in each individual study. Methods to be developed within the *metaSURV* can also be used to define safe intervals after negative screen-test results and to develop longitudinal criteria for new screening tests other than hrHPV DNA assays for which the cross-sectional criteria of the Meier guideline are not accepted. Although graphical displays and derived data are approximate and depend on the censoring rule, the methods tentatively outlined above, allow a quick and rather robust interpretation of the longitudinal safety associated with different screening policies.

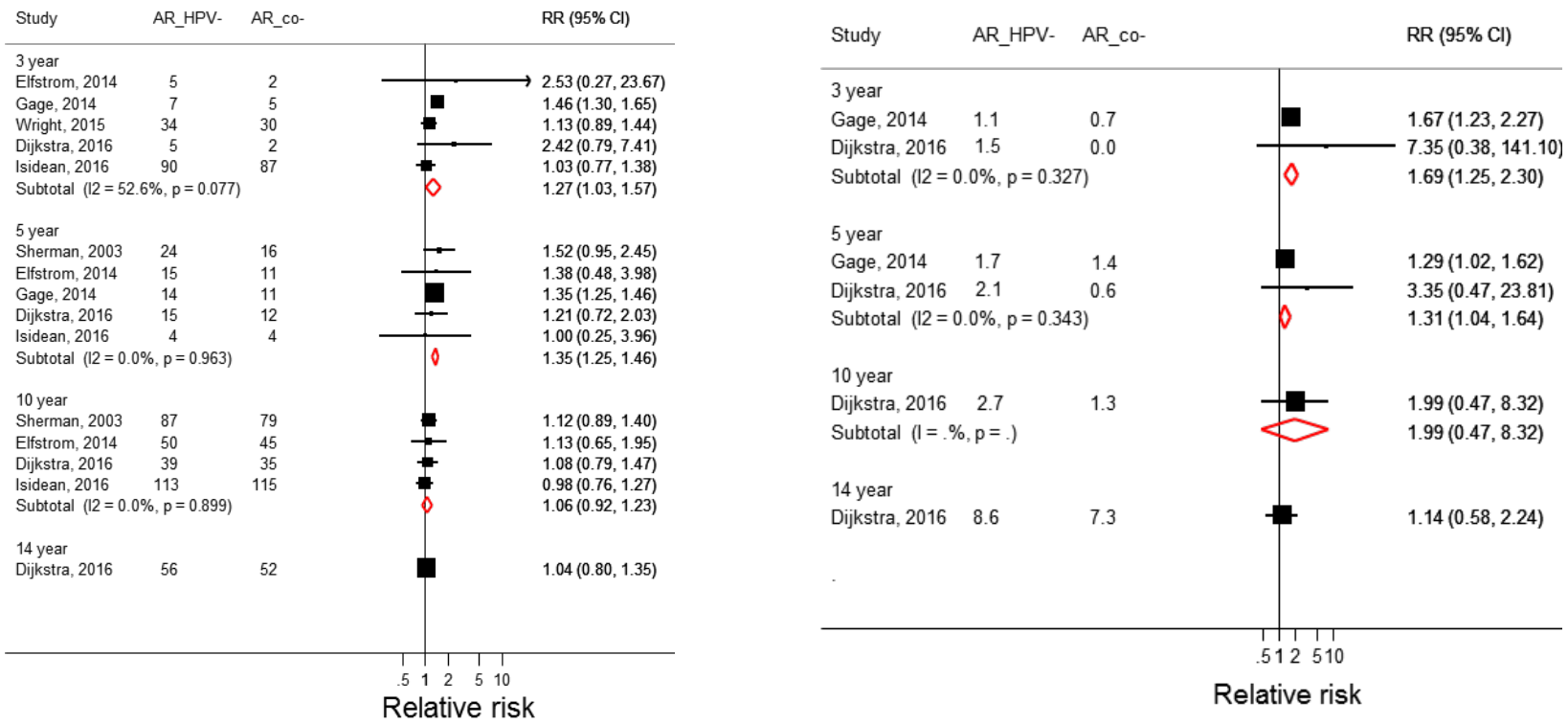


Figure 49. Relative risk of CIN3+ (left) or invasive cervical cancer (right) after a negative HPV test vs after a negative co-test, observed after 3, 5, 10 and 14 years of observation. The columns AR_HPV- and AR_co- contain the absolute cumulative risk per 10,000 screened women.

Note: the relative risk of cancer after 3 years in Dijkstra et al, 2016 required a continuity correction because of zero cases after co-testing, to compute the relative risk (+0.5 to each cell of the 2x2 table).

12.6.3.Cervical cancer screening of vaccinated cohorts

All data pooled in the systematic reviews, presented in this PhD manuscript, concerned women who were not vaccinated against HPV. Over the last decade, several countries around the world have included HPV vaccination in their national routine immunisation programmes, which typically target girls in the age range 11-15 years[335] In addition, some countries have included opportunistic vaccination or catch-vaccination up to the age of 18-26 years. The earliest vaccinated cohorts, who belonged to the catch-up vaccination group, are currently reaching the target age for cervical cancer screening. Surveillance data have shown an age gradient in vaccination effectiveness. Subjects, vaccinated at an older age, (let us say at 18 years or older and already exposed to prior HPV infection) show a more moderate protection against cervical pre-cancer due to the vaccine types. Levels of protection are typically in range observed in the intention-to-treat analyses of the vaccination trials[336]. Females vaccinated at younger age, tend to show an effectiveness in the range observed in the per-protocol analyses of the vaccine trials[144,336]. Future screening policies should cope with this changing landscape and take into account that the risk of cervical pre-cancer and invasive cancer will drop progressively. Scotland (starting cytology-based screening at the age of 20 years) is the first European country that was able to observe the impact of vaccination on predictive values of cytological screening. As expected, the positive predictive value of HSIL cytology for detecting underlying CIN2+ lesions decreased (-32%, 95% CI 14-46%) and, consequently, the number of colposcopy referrals needed to find one histologically confirmed CIN2+ increased[337]. It has been anticipated that due to the decrease in the number of HSIL lesions and the relative dominance of LSIL and ASC-US lesions, performance of microscopic detection of cytological abnormalities would become challenging and considerably less motivating in vaccinated cohorts. On other hand, automated HPV testing could be more easily adapted by considering higher viral load cut-offs to maintain positive predictive power[338].

Recent epidemiological review and modelling work suggest following adaptations of cervical cancer screening policies for vaccinated cohorts: start screening at later age (30 or older), screening only with HPV tests, longer intervals than five years for HPV negative women and use of more specific tests[339].

12.6.4. Cost-effectiveness of HPV-based screening

A recent Belgian health-technology assessment compared the cost-effectiveness (CE) of screening by cytology with screening by HPV testing and concluded that by switching from cytology screening at 3-year intervals to HPV-based screening at 5-year intervals was cost-effective and even cost-saving[111]. It was estimated that for 100,000 women entering HPV-based screening at the age 30 followed for the rest of their life during 74 years[k], 240 cases of cervical cancer, 178 deaths from cervical cancer were avoided and 1,618 life years were gained for a discounted cost that was almost 6 million € less, compared to screening with cytology. Basic costs for HPV testing were estimated at 35€/test. Since costs for HPV tests have dropped substantially since recent years and this evolution appears to continue in the years to come, conclusions are expected to become

[k] Follow-up up to the age of 104 years where all women were assumed to have died.

even more convincing in the future. The findings of this Belgian HTA are in line with other cost-effect evaluations from other high income countries[340,341].

Three recent economic analyses were published that assessed costs and effects of cervical cancer screening using strategies with HPV testing alone and co-testing in the USA, Australia and New Zealand[342-344]. All the studies concluded consistently that the incremental costs for the rather small gain in quality-adjusted life-years associated with co-testing exceeded the usual willingness-to-pay thresholds. The results did not change when the evaluated screening strategies were simulated on cohorts vaccinated against HPV.

12.7. HPV testing on self-samples

An additional advantage of using HPV testing is that, contrary to cytology[309], it can be done on a vaginal sample collected by the woman herself. Offering women a device for self-sampling can increase the population coverage by reaching those who are reluctant to participate in the regular screening program that require clinic-based visits and pelvic exams[112,345]**.** A recent meta-analysis of randomised trials, concluded that strategies, including the offer of a self-sample, were more effective in reaching non- or under-screened women than sending reminders to have a cervical specimen taken by a health professional[112].
Self-sampling has its interest in primary screening. However, evaluation of clinical accuracy in screening requires large sample sizes because of the low prevalence of high-grade CIN in the general population. The finding, derived from our meta-analysis[308], that the relative sensitivity and specificity of HPV testing on self- versus clinician-taken samples were similar for screening and follow-up settings, offers interesting opportunities for efficient study designs to evaluate the relative accuracy of combinations of self-sampling devices and HPV assays in a colposcopy clinic.

HPV testing is as accurate on self-samples as on clinician-taken samples under the condition that clinically validated PCR-based assays are used[309]. Our meta-analysis[308] did not reveal effects of the device used for self-sampling. However, included studies most often did not compare self-samplers within the same study. A recent Germany study, showed higher viral loads assessed with the Abbott RealTime High Risk HPV test, in samples taken with the Evalyn-Brush (Rovers Medical Devices, the Netherlands) compared to the Qvintip (Aprovix, Sweden) but the clinical relative accuracy for CIN2+ or CIN3+ (self vs clinician samples) was similar for both self-devices[346].

The lower clinical sensitivity of HPV testing with HC2, Cervista or APTIMA on self-samples can be explained by lower loads of high-risk HPV DNA from virus particles in the vagina which are beneath the detection threshold of respective assays, but which can be detected by analytically more sensitive PCR tests[297,310]. In addition, the lower specificity of HC2, Cervista may be ascribed partly to increased presence of cross-reacting low-risk HPV types in the vagina but also to presence of high-risk HPV particles in the vagina which are not associated with CIN2+ lesions[310].

Reflex cytology is recommended to triage HPV-positive women in an HPV-based screening programme, where liquid-based specimens are taken by health professionals. (see Results section 10.3). The residual material remnant after HPV testing can be used for cytological processing. Since cytology triage is sub-optimal on self-samples, an adequate molecular reflex test to triage HPV-positive cases would be very useful. Candidate triage methods are the currently available genotyping for HPV16/18 [200], which accounts for ~70% of the cancer risk, and, in the future maybe also methylation markers of certain viral or human genes[347,348] which both are associated with progressing infections.

12.8. Limitations of meta-analyses

The meta-analyses assessing triage of women with minor cytology, follow-up after treatment of cervical precancer and HPV testing of self-samples pooled data on test accuracy and only rarely included randomised trials. Diagnostic test accuracy studies are observational in kind and therefor generate evidence of low to moderate quality[349].

A general limitation inherent to meta-analyses of aggregated data extracted from published data is the limited number of potentially influential covariates that could be accounted for. We were unable to perform subgroup meta-analyses or to conduct meta-regressions that incorporated influential factors such as age, screening history, presence of risk factors. To address this limitation, individual patient data meta-analysis should be set up. The efficacy of cytology versus HPV-based screening could be assessed from large randomised population trial. In addition, the individual records with follow-up over several years of four RCTs could be pooled allowing assessment of age effects and safe screening intervals.

Meta-analyses sometimes give the impression to yield precise pooled estimates, even if random effect models are applied. However, as stated in the Methods, readers should not only focus on the pooled overall measures but pay also attention to the spread of observations. In our meta-analyses, we showed the inter-study variability in forest plots and hSROC plots. We always quantified the inter-study heterogeneity and tried to explain the sources of heterogeneity. However this was not always possible by lack of uniformly reported covariate data.

Meta-analysis increase the statistical power to find differences which in individual studies may lack. However, where insignificant findings from individual poorly designed studies, protect against over-interpretation, meta-analyses can accumulate bias and yield statistically significant but spurious findings. Therefor careful judging of study quality using standard tools as QUADAS and the Cochrane tool for assessment of bias in RCTs is a crucial part of a good meta-analysis. In our meta-analysis of the accuracy of HPV testing on self-samples, we integrated all the QUADAS items in a meta-regression. We could not identify major design issues influencing the accuracy. However, the judgement

of the QUADAS items was not always easy by lack of reported information. A typical example is partial verification bias, which cannot be proven in studies, where authors state that all subjects received colposcopy and biopsy, and omit information on those who were not followed.

More suggestions to improve the quality of future meta-analyses are included in Chapter 13.

13. Conclusions and propositions for further research

13.1. Conclusions

13.1.1. Triage of women with minor abnormal cervical cytology

1. It has been consistently demonstrated that triage of women with ASC-US with a hrHPV DNA test is more sensitive and similarly specific compared to repeat cytology to detect underlying cervical CIN2+ and CIN3+.The meta-analyses included in this PhD thesis were recognised as evidence base for practice recommendations in Europe, and other continents and has resulted in decision making in Belgium to reimburse hrHPV testing for women with ASC-US.

2. Triage of women with LSIL with a hrHPV DNA test is more sensitive but substantially less specific than repeat cytology to detect underlying CIN2+ or CIN3+. The posttest probability of CIN3+ is hardly different from the pretest probability, indicating poor utility.

3. Triage of ASC-US and LSIL with a test, targeting RNA of all hrHPV types, is as sensitive but more specific than HC2 or other validated hrHPV DNA assays.

4. Triage of ASC-US and LSIL with a test targeting HPV RNA of only five hrHPV types (HPV16, HPV18, HPV31, HPV33, HPV45) is significantly less sensitive than assays targeting DNA or RNA of 13/14 hrHPV types.

5. Overexpression of p16 or p16 & Ki67, assessed by cyto-immunochemistry, is a strong indicator for presence of CIN2+ or CIN3+. Nevertheless the sensitivity is slightly lower than hrHPV assays.

6. Genotyping for HPV16 or HPV16/18 as sole test is insufficiently sensitive to triage women with ASC-US of LSIL. However, when used to triage hrHPV+ women with ASC-US or LSIL, a distinction can be made between HPV16/18+ women who should be referred to colposcopy and HPV16/18- negative women who can be followed-up conservatively (need a second visit for retesting).

7. An optimal triage test or algorithm including a combination of triage tests, upgrades the probability of pre-cancer (CIN2+ of CIN3+) when the triage result is positive to a level where referral for diagnostic work-up is justified and downgrades the probability of pre-cancer below a threshold where safe release to the routine screening is considered as safe. Each programme can define its own target decision thresholds. The Pretest-Posttest-Probability plot provides an easily understood graphical display of the clinical utility of a triage test.

13.1.2. hrHPV testing as a test of cure after treatment of cervical pre-cancer

1. Women treated for cervical pre-cancer (CIN2, CIN3 or AIS) have a significantly increased risk of residual or recurrent pre-cancer or cancer compared to the general population. Therefore, careful post-treatment surveillance is required.

2. Post-treatment follow-up with hrHPV testing is more sensitive and almost as specific compared to cytology to predict residual or recurrent high-grade CIN.

3. Post-treatment co-testing with a hrHPV test and cytology is more sensitive but less specific compared to HPV testing only.

4. A safe algorithm, could the following: co-testing with a hrHPV test and cytology at 6 months after treatment followed by another co-test at 18 months if the tests at 6 months were negative with referral to colposcopy if at least one test was positive and referral to the normal screening schedule when all tests were negative. Women being twice co-test negative have a future risk of CIN3+ which is similar to women with a negative cytological screening test.

13.1.3. hrHPV testing in primary cervical cancer screening

1. Cervical cancer screening using a hrHPV assay is more sensitive but less specific than screening with cytology to detect underlying or incipient cervical cancer precursors. However, the observation of the higher sensitivity on its own does not provide sufficient evidence to recommend HPV-based cervical cancer screening as first choice test given the spontaneous regression of most cases of pre-cancer. Indeed, from cross-sectional data, we cannot exclude that HPV-based screening just would detect only more regressive CIN2+ than cytology. Therefore longitudinal evidence is needed showing lower risk of CIN3+ or cancer after a negative hrHPV test compared to after a negative Pap smear. By lack of this longitudinal evidence, HPV-based screening was not yet recommended in the 2nd edition of the European Guidelines for Quality Assurance in Cervical Cancer Screening.

2. In the meanwhile, randomised trials have demonstrated a lower cumulative incidence not only of CIN3+ but also of invasive cervical cancer among women with a baseline negative hrHPV DNA result. This observed fact provides high-level evidence justifying a strong recommendation to switch to HPV-based cervical cancer screening.

3. HPV-based screening should not start before the age of thirty, given the high prevalence of transient hrHPV infections in adult women before that age. Before the age of thirty, cytology-based screening is preferred.

4. Given the higher sensitivity and longer lead time, HPV-based screening can be offered at longer intervals compared to cytology-based screening. The interval should not be shorter than five years.

5. No clear evidence is available regarding the question whether HPV-based screening can be stopped at an earlier age than cytology-based screening, which is in the range of 60-65 years in European guidelines.

6. In randomised trials, the Hybrid Capture II assay and the GP5+/6+-EIA PCR were used. Both assays detect DNA of hrHPV types. Therefore these assays are accepted for use in cervical cancer screening. Other hrHPV DNA which demonstrate good reproducibility and non-inferior sensitivity and specificity may be accepted for primary cervical cancer screening. A list of hrHPV DNA assays that fulfil these criteria is provided in sub-chapter 10.6. Assays targeting other molecules (RNA, methylation, protein markers) require demonstration of longitudinal safety (low risk of CIN3+ or cancer after negative testing over a given period).

13.1.4. HPV testing on self-samples

1. Sending devices for vaginal self-sampling at home is more effective to raise cervical cancer screening coverage than sending reminder letters to have a Pap smear taken by a clinician.

2. hrHPV DNA testing on self-samples is as accurate as on clinician-taken samples under the condition that a clinically validated PCR-based hrHPV DNA test is used. hrHPV DNA tests, that are based on a principle of signal amplification (HC2, Cervista), are less sensitive and also less specific on a self-sample than on a clinician-taken sample. hrHPV RNA testing is less sensitive on self-samples but as specific on self- compared to on clinician-taken samples.

3. No effect of the self-sampling device on the accuracy on hrHPV testing on self-samples to detect high-grade CIN could be identified. However, since the studies included in the meta-analyses essentially compared self- with clinician sampling no formal conclusion on the best choice of self-sampling device can be formulated.

13.2. Propositions for further research

Besides synthesising the currently available evidence regarding a clinical question, systematic reviews also identify gaps in knowledge, suggesting need for new studies. Below, we provide suggestions for further research that came up through the meta-analyses on the four applications of HPV testing, discussed in the previous chapters.

13.2.1. Triage of women with minor abnormal cervical cytology

The massive amount of studies and different tests that were included in several meta-analyses have generated a unique database which can be used as template for new statistical procedures, such as network-meta-analyses, pretest-pretest probability plots.

At the Belgian Cancer Centre, a new Bayesian procedure in R has been worked out, where a multitude of triage tests can be evaluated at the same time[350-352]. The data will also be used as a substrate for an extended version of the METADAS SAS macro, allowing the inclusion of more than one co-variate[138]. The METADAS SAS procedure, based on the bivariate or hsROC models, currently recommended in diagnostic test accuracy Cochrane reviews, accepts only one covariate.
The PPP plot, which was introduced in one of our recent publications[145], is a useful graphical tool to demonstrate the utility of a diagnostic test in a given situation,. More research needs to be done to incorporate statistical inference regarding the joint variability in accuracy and back ground prevalence[353].

Triage of screen-positive subjects offers an efficient setting to evaluate new biomarkers, that could be considered, in a later stage, in primary screening. Because of the higher background prevalence of cervical pre-cancer, verification of all subjects, using study designs recommended for good diagnostic accuracy research with a valid reference standard, is ethically acceptable and economically feasible in a triage environment.

13.2.2. Follow-up of women treated for cervical pre-cancer

A very recent review published after the completion of this PhD manuscript, concerned the prediction of cure or failure of cervical pre-cancer therapy and revealed that positive section margin are associated with recurrent high-grade CIN[354]. In another context, we also assessed the obstetrical harm associated with excision or ablation of intraepithelial cervical lesions[89,103,244]. Conflicting recommendations resulted from these reviews: removal of sufficient neoplastic tissue to assure cure versus prudence to avoid future obstetrical harm. Up hereto, no studies assessed both outcomes (oncological versus obstetrical safety). Therefore it is proposed to countries with good registries to target research to incorporate both outcomes. This research proposition was recently granted and will be addressed in a nationwide Norwegian study.

13.2.3. Primary cervical cancer screening

The VALGENT protocol offers a comprehensive research forum to validate new HPV assays which can potentially be included in primary cervical cancer screening[251]. The VALGENT-3 and -4 studies are currently running and involve laboratories and HPV assay manufacturers from Europe, USA, China and Australia. This test validation network appears to be robust and contributes in generating evidence regarding assays that fulfil requirements for screening. The continuously updated lists of validated assays facilitate national and international decision making regarding the choice of a particular HPV test in screening programmes.
The cross-sectional design of VALGENT limits the scope to validation of DNA-based HPV assays. Internationally accepted longitudinal criteria need to be developed which would allow use of assays targeting other molecules potentially usable in screening (HPV RNA, protein markers, methylation markers, ...). A proposal based on cumulative risk over time of CIN3 or cervical cancer after a negative test was presented at the 30th International Conference of the Papillomavirus Society in Cape Town (March, 2017).

Subsequent to that conference, the Unit of Cancer Epidemiology of the Belgian Cancer Centre was appointed to prepare the future update of the current guidelines defining the minimal requirements of HPV assays acceptable for cervical cancer screening and to develop longitudinal performance criteria for new tests targeting other molecules than HPV DNA.

Since more countries have switched or consider to switch to HPV-based cervical cancer screening, our focus for future meta-analyses will move towards triage of HPV-positive women.

Another priority for research is the adaptation of screening policies for future generations of young women, where HPV vaccination coverage will increase progressively over time.

As indicated in § 12.6.1, the existence of hrHPV negative cervical tumours, invites for research to assess the burden of cancers that would be missed with screening using only HPV tests. Systematic audits of HPV negative cervical cancers and matched HPV positive cancers involving panels of multiple HPV tests targeting different viral genes and NGS assays applied on the biopsies as well as on archived prior cervical cell specimen may offer clues for better screening tests for the future.

13.2.4. HPV testing on self-samples

Our meta-analysis on the accuracy of HPV testing on self-samples revealed important HPV test effects but was not able to assess the influence of the devices used for self-sampling[308,309]. The VALHUDES[1] protocol is designed to evaluate the clinical accuracy of HPV assay – self sampler combinations as well as HPV testing on urine samples. This protocol enrols 500 women referred to colposcopy, who will take a urine sample with the Colli-Pee device (Novosanis, Wijnegem, Belgium), a vaginal self-sample and a cervical sample taken by a gynaecologist.

A meta-analysis, also conducted at the Belgian Cancer Centre, showed that sending self-samplers to women who do not participate in the regular screening programme is more effective than sending reminders to have a sample taken by a clinician[112]. However, the rate of response to home mailing of self-sampler devices still is limited and essentially reaches the better socio-economic classes among the non-screened women. General practitioners may identify women without screening records who consult for whatever reason. GPs could take this opportunity to offer a self-sampling device. A recent small trial conducted in a GP practice in Brussels showed promising results which merit to be assessed in a large scale study[355].
In general, more research is needed to evaluate strategies which may optimise participation in preventive health interventions using rigid randomised study designs.

[1] VALHUDES: VALidation of Human papillomavirus assays and collection DEvices for HPV testing on Self-samples and urine samples.

13.3. Internationalisation of health technology assessment

We observe nowadays that decision making regarding introduction on new technologies in clinical practice is preceded by systematic review work. Performing high-quality reviews requires skills in retrieving and synthesizing data and sufficient knowledge of the pathology and its clinical aspects[356]. The combination of both competencies are rarely available within one centre. Health technology assessments of new screening and vaccination strategies are often conducted simultaneously in several countries and institutions. This results in multiplication of resources, dilution of competencies and sometimes yields contradictory findings, generating confusion among stakeholders, health professionals and the general public. International coordination is needed involving specialists, skilled in health-technology assessment, HPV epidemiology and clinical experts, allowing for balanced interests. We therefore advocate to internationalise more the efforts to conduct new systematic reviews. We recommend decision makers not to repeat systematic reviews already done and to focus on cost-effectiveness and implementation questions. The creation of a Cochrane group dedicated to prevention and treatment of HPV-related disease could provide an internationally recognised forum that could provide services for local or international needs.

It is the ambition of the Cancer Centre of the Scientific Institute of Public Health to contribute in keeping the evidence on primary and secondary prevention of cervical cancer updated and to extend the work towards prevention and treatment of other HPV-related pathologies as well. Triage of minor cytological lesions will become a less important topic because of the general move towards HPV-based screening. Priorities for further meta-analytical work are: 1) triage of HPV-positive women, including genotyping, methylation, cell-cycle protein markers; 2) management options to follow-up women after treatment of cervical pre-cancer, including assessment of the section margins and post-treatment testing with biomarkers; 3) individual patient data meta-analysis on the correlation between the dimensions of excised tissue and subsequent risk of preterm delivery; 4) diagnosis/prognosis of anal cancer and of HPV-related head-and neck-cancers; 5) Cochrane reviews on safety and efficacy of HPV vaccination against HPV infection and associated cervical, vulvar, vaginal, anal and penile (pre-)cancer. Review work on all these topics is on-going. International collaborations, networking and training of young dedicated researchers in the field of evidence assessment must be a continuous point of attention.

14. Acknowledgements

I thank my three promoters Prof Dr Marleen Temmerman, Prof Dr Steven Weyers and Prof Dr John-Paul Bogers, for their continued motivation and advice. I am also grateful for my Directors Prof Dr Herman Van Oyen, Dr Viviane Van Casteren and Dr Marc van den Bulcke for the opportunity to work for two decades at the Scientific Institute of Public Health in Brussels and the opportunities to develop my skills in systematic reviews. In particular, I would like to mention my colleagues Cindy Simoens, Jolien Roelens, Freija Verdoodt, Victoria Nyaga, Lan Xu, Eliana Peeters and Paula Busto who all have contributed to our systematic review work.

The production of all the systematic reviews, meta-analyses, Cochrane reviews, health technology reports and practice guidelines handled in the PhD manuscript have required a lot of time which I could not spent with my family. I therefor thank my family, in particular my wife Julie and our children Daniella, Hans and Katrien for their support and I must apologize at the same time for my lack of availability.

15. Annexes

15.1. Terminology used for reporting of cervical cytological or histological findings

Table 30. Conversion table for different cytological and/or histological classification systems used to grade cervical (pre-)cancerous lesions of the uterine cervix. Adapted from the 2nd edition of the European Guidelines for Quality Assurance in Cervical Cancer Screening[221,223,357].

<table>
<tr><th>Papanicolaou</th><th>WHO</th><th>CIN (Richart, 1973)[7]</th><th>TBS 1991 (Luff, 1992)[358]</th><th>TBS 2001 (Solomon, 2002)[8]</th><th>LAST (Darragh, 2012)[359]</th></tr>
<tr><td>Cytology</td><td>Histology</td><td>Histology</td><td>Cytology</td><td>Cytology</td><td>Histology</td></tr>
<tr><td>I</td><td colspan="3">Normal</td><td rowspan="2">NILM</td><td rowspan="2"></td></tr>
<tr><td rowspan="4">II</td><td colspan="2" rowspan="3">Atypia</td><td>Infection, reactive repair</td></tr>
<tr><td rowspan="2">ASCUS</td><td>ASC-US</td><td></td></tr>
<tr><td>ASC-H</td><td></td></tr>
<tr><td colspan="2">Atypical glandular cells</td><td>AGUS</td><td>AGC</td><td></td></tr>
<tr><td rowspan="5">III</td><td rowspan="2">Mild dysplasia</td><td>Condyloma</td><td rowspan="2">LSIL</td><td rowspan="2">LSIL</td><td rowspan="2">LSIL

(including p16-CIN2, p16- unclear CIN)</td></tr>
<tr><td>CIN I</td></tr>
<tr><td rowspan="2">Moderate dysplasia</td><td rowspan="2">CIN II</td><td rowspan="3">HSIL</td><td rowspan="3">HSIL</td></tr>
<tr><td rowspan="3">IV</td><td rowspan="2">Severe dysplasia</td><td rowspan="2">CIN III</td><td rowspan="2">HSIL (including p16+ CIN2, p16+ unclear CIN)</td></tr>
<tr><td>CIS</td></tr>
<tr><td>AIS</td><td>CGIN</td><td>AGUS</td><td>AIS</td><td></td></tr>
<tr><td>V</td><td colspan="3" rowspan="2">Invasive carcinoma (micro-invasive, stages I-IV) (squamous cell carcinoma, adeno-carcinoma, adenosquamous carcinoma)</td><td>SISCCA</td></tr>
<tr><td></td></tr>
</table>

AGC: atypical glandular cells; AGUS: atypical glandular cells of undetermined significance; ASC-H: atypical squamous cells cannot exclude HSIL; AIS: adenocarcinoma in situ; ASC-US: atypical squamous cells of undetermined significance; CGIN: cervical glandular intraepithelial neoplasia; CIN: cervical intraepithelial neoplasia; CIS: carcinoma in situ; HSIL: high -grade squamous intraepithelial lesion; LAST: Lower Anogenital Squamous Terminology Standardization; LSIL: low-grade squamous intraepithelial lesion; NILM: Negative for epithelial abnormality; SSSCCA: superficially invasive squamous cell carcinoma ;TBS: the Bethesda System;WHO: World Health Organisation.

15.2. Virological test characteristics of HPV assays evaluated in the PhD project

Table 31. Virological test characteristics of hrHPV DNA and RNA tests evaluated in the PhD project.

hrHPV assay [Ref]	Target, *genes*	Type of amplification	Identification of HPV	Separate genotyping	Internal control for human genes	Quantitative signal correlated with viral load
Comparator tests						
Hybrid Capture-2 (HC2) [360]	DNA, *several viral genes, un-defined*	SA Hybridisation with coated RNA.	Antibody against DNA-RNA hybrids	None. B probe detects in bulk 13 hr types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68). A probe detects in bulk 5 lr types: (6,11,42,43,44).	No	Intensity of light emission expressed as RLU (relative to 1pg/mL of HPV17 DNA).
GP5+/6+ PCR-EIA [361]	DNA, *L1 (conserved 140 bp sequence)*	TA, PCR with GP5+/6+ primers.	EIA recognition of amplicons hybridized with 14 oligo-nucleotids.	None. Detects in bulk 14 hr types: same as HC2 plus HPV66.	No	None

hrHPV assay [Ref]	Target, *genes*	Type of amplification	Identification of HPV	Separate genotyping	Internal control for human genes	Quantitative signal correlated with viral load
HPV assays, evaluated for triage of minor abnormal cervical cytology or potentially usable in cervical cancer screening						
Abbott RealTime High Risk HPV test [253-255]	DNA, *L1*	TA, by RT PCR with modified GP5+/6+ primer mix.	4 channels for detection of fluorescent signals.	16,18 & 12 other hr types (see GP5+/6+ PCR-EIA).	*β-globin*	Cycle number
Anyplex II HPV HR [273,274]	DNA, *L1*	TA, 2 multiplex reactions on the CFX96 RT PCR	Pitcher/ catcher system.	Separate identification of 14 hr types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68).	*β-globin*	Signal: +, ++, +++
Anyplex II HPV28 [362,363]	DNA, *L1*	See Anyplex II HPV HR	See Anyplex II HPV HR.	Separate identification of 28 types: 14 hr types (16,18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 66, 68); 6 phr types (26, 53, 69, 70, 73, 82); 8 lr types (6, 11, 40, 42, 43, 44, 54, 61).	*β-globin*	See Anyplex II HPV HR
APTIMA HPV assay[252,364]	RNA, *E6/E7*	TA by transcription-mediated amplification	Hybridization Protection Assay.	No (a prototype identifying HPV16, 18-45 exists)	No. Assay includes internal controls for non-infectious RNA and DNA.	None.

hrHPV assay [Ref]	Target, *genes*	Type of amplification	Identification of HPV	Separate genotyping	Internal control for human genes	Quantitative signal correlated with viral load
*care*HPV Test [266]	DNA, *several viral genes, un-defined*	SA, see HC2	see HC2	See HC2	No	See HC2
Cervista HPV HR Test [257,258]	DNA, *L1/E6/E7*	SA using Invader chemistry [365].	binding of 5'flaps to FRET oligonucleotides, generating fluorescence signal.	14 hr types. Separate typing of 16, 18 available as a reflex test.	human *histone 2*	FAM fluorescence signal.
cobas 4800 HPV test [259,260]	DNA, *L1*	TA by RT PCR	probes with 4 different reporter dyes	Separate identification of types16,18 and bulk of 12 other hr types (31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68).	*β-globin*	None
EUROArray HPV [366]	DNA, *E6/E7*	TA multiplex qPCR	Oligo-nucleotide microarray system.	Separate identification of 30 types: 14 hr types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68); 4phr types (26, 53, 73, 82); 12 lr or ur types (6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81, 89).	*Hsp70*	None

hrHPV assay [Ref]	Target, *genes*	Type of amplification	Identification of HPV	Separate genotyping	Internal control for human genes	Quantitative signal correlated with viral load
GP5+/6+-LMNX [265]	DNA, *L1*	TA, PCR with GP5+/6+ primers	Luminex Genotyping kit	Separate typing of hr HPV types 16,18, 31,33,35,39,45,51,52,56,58,59, 66,68; Detects also phr types:26,53, 73, 82.	Human DNA fragment located on chromosome 14	None
HPV-Risk assay [262]	DNA, *E7*	TA by multiplex real-time PCR targeting a ~150-bp fragment.	Hydrolysis probes with 3 spectrally unique fluorescent dyes	HPV16, 18 and 13 other hr types (see HC2 plus 66 & 67)	*β-globin*	CT value
INNO-LiPA HPV Genotyping Extra II [367]	DNA, *L1*	TA by RT PCR with SPF10 primers targeting a sequence of 65bp.	Line blot hybridisation.	Separate identification of 28 different HPV types, 14 hr types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68); 5 phr types (26,53,73, 70,82); 6 lr types (6, 11, 40, 43, 44, 54); 3 types with undetermined risk (69, 71, 74).	*HLA-DPB1 gene*	None

hrHPV assay [Ref]	Target, *genes*	Type of amplification	Identification of HPV	Separate genotyping	Internal control for human genes	Quantitative signal correlated with viral load
Linear Array [368]	DNA, *L1*	TA by PCR with PGMY09/11 primers.	Reverse line blot hybridization.	Separate typing of 37 types: 14 hr types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68); 4 phr types (26, 53, 73, 82; 10) lr types (6, 11, 40, 42, 54, 61, 70, 72, 81), CP6108; and 9 types with undetermined risk (55, 62, 64, 67, 69, 71, 83, 84, IS39).	*β-globin*	None
MALDI-TOF [267,369]	DNA, *L1*	TA by 1ary PCR with ts GP5+/6+ primers, followed by mass extension reaction with a single primer of distinct mass that is also ts.	Mass-spectrometry	Separate typing of 14 hr HPV types 16,18, 31,33,35,39,45,51,52,56,58,59, 66,68.	*β-globin*	None
Onclarity HPV assay [256,271,370]	DNA, E6/E7	TA by RT PCR.	Fluorescent probes for 4 separate channels.	Separate typing of 6 hr types 16,18,31,45, 51,52; typing of groups 33-58; 56-59-66; 35-39-68.	*β-globin*	None

hrHPV assay [Ref]	Target, *genes*	Type of amplification	Identification of HPV	Separate genotyping	Internal control for human genes	Quantitative signal correlated with viral load
PapilloCheck HPV-screening test [263]	DNA, *E1 (*350-bp sequence)	TA by RT PCR	Ts oligonucleotide probes immobilized on a DNA chip.	Separate typing of 23 types: 14 hr types 16,18,31,33,35,39,45,51,52,53, 56,58,59,66, 68; 3 phr types 70,73,82; 6 lr types 6,11,40,42,43,44.	*ADAT1*	None
Pretect HPV-Proofer[268,371]	RNA, *E7/E7*	TA by RNA-based RT nucleic acid sequence-based amplification.	Molecular beacon hybridisation	Separate typing of 16,18,31,33, 45.	U1 small nuclear ribonucleo-protein-specific mRNA.	None
Riatol qPCR(E6/E7) [261]	DNA, *E6/E7*	TA by RT qPCR	Ts RT qPCR.	Separate typing of 17 types: 14 hr types (16,18,31,33,35,39,45,51,52,56, 58,59,66, 68); one phr type (53) and lr types (6, 11).	*β-globin*	HPV copies/cell
Xpert HPV [269,269,275,372]	DNA, *E6/E7*	TA RT PCR	14 hrHPV types detected in 5 fluorescent channels.	Separate identification of HPV16, and four groups of hr types (18-45, 31-33-52-58, 51-59, 39-56-66-68).	*hydroxymethylbilane synthase (HMBS)*	CT value

CT: cycle threshold; EIA: enzyme-immuno assay; PCR: polymerase chain reaction; hr: high-risk; lr: low-risk; PCR: polymerase chain reaction; phr: possible high-risk; qPCR: quantitative PCR; RT: real time; SA: signal amplification; TA: Target amplification; ur: undetermined risk; ts: type-specific.

15.3. Search strings

Box 2. Embase search string used for literature retrieval regarding the accuracy of HPV tests used to triage women with ASC-US or LSIL.

'cervix'/exp OR cervic* AND ('cancer'/exp OR cancer OR 'carcinoma'/exp OR carcinoma OR 'neoplasia'/exp OR neoplasia OR 'dysplasia'/exp OR dysplasia OR cin OR sil) OR ('cervix'/exp OR cervix AND ('neoplasm'/exp OR neoplasm)) AND ('hpv'/exp OR hpv OR 'wart virus'/exp OR 'wart virus' OR ('hpv'/exp OR hpv AND ('dna'/exp OR dna)) OR ('hpv'/exp OR hpv AND viral) OR 'human papillomavirus'/exp OR 'human papillomavirus' OR 'hybrid'/exp OR hybrid AND capture OR 'hc2 assay' OR hc2 OR 'hc2' OR hcii OR 'hc ii') AND ('triage'/exp OR triage OR 'management'/exp OR management OR 'follow up'/exp OR 'follow up')

Box 3. Medline search string used for literature retrieval regarding the accuracy of HPV tests used in follow-up after treatment of CIN.

((cancer OR carcinoma OR dysplas* OR neoplas* OR CIN OR SIL) AND (cervix OR cervical) OR vaginal smears[MeSH] OR Cervix Neoplasm [Mesh])) AND (treatment OR conisation OR conization OR LLETZ OR leep OR cryotherapy) AND (HPV OR human papillomavirus OR papillomavirus)

Box 4. Medline search string used for literature retrieval regarding the accuracy of HPV tests in cervical cancer screening.

((Uterine Cervical Neoplasms [MeSH Terms] OR Uterine Cervical Dysplasia [MeSH Terms] OR Cervical Intraepithelial Neoplasia [MeSH Terms] 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 (papillomaviridae [MeSH:NoExp] OR alphapapillomavirus [MeSH Terms] OR "DNA, viral" [MeSH Terms] OR Papillomavirus Infections [MeSH Terms] OR Tumor Virus Infections [MeSH Terms] OR "Cervix Uteri/virology" [MeSH Terms] 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

Box 4 (continued)
(Vaginal smears [MeSH Terms] OR Cytodiagnosis [MeSH Terms] OR Cell Transformation, Viral [MeSH Terms] OR Cytopathogenic Effect, Viral [MeSH Terms] 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 [tw] OR screen* OR "Early Detection of Cancer"[Mesh])

Box 5. Medline search string used for literature retrieval regarding the accuracy of HPV tests on self-samples.

#1: Cervix OR cervico* OR cervica*
#2: Cancer OR carcinoma OR neoplas* OR dysplas* OR CIN[tw] OR CINII*[tw] OR CIN2*[tw] OR CINIII*[tw] OR CIN3[tw] OR SIL[tw] OR SIL OR HSIL[tw] OR H-SIL OR LSIL[tw] OR L-SIL OR OR ‘‘low grade’’ OR low-grade OR mild OR equivocal OR borderline.
#3: #1 AND #2.
#4: HPV OR "Human Papillomavirus DNA Tests"[Mesh] OR ‘‘human papillomavirus’’ OR papillomavir* OR viral OR virus
#5: self-collection OR “self collection” OR self-sampling OR self-collect* OR self-sampl* OR self OR "Self-Examination"[Mesh]
#6: #4 AND #5
#7: #3 AND #6
#8: Publication Date to May 25, 2013
#9: #7 AND #8

Box 6. Embase search string used for literature retrieval regarding the accuracy of HPV tests on self-samples.

#1: 'cervix'/exp OR cervix OR cervico* OR cervica*
#2: 'cancer'/exp OR cancer OR 'carcinoma'/exp OR carcinoma OR neoplas* OR dysplas* OR cin OR 'cin2' OR 'cin3' OR sil OR hsil OR h+sil OR lsil OR l+sil OR 'low grade' OR low+grade OR mild OR equivocal OR 'borderline'/exp OR borderline
#3: 'hpv'/exp OR hpv OR 'human papillomavirus'/exp OR 'human papillomavirus' OR papillomavir* OR viral OR 'virus'/exp OR virus
#4: self+collection OR 'self collection' OR self+sampling OR 'self-sampling' OR self+collect* OR self+sampl* OR 'self'/exp OR self
#5: #1 AND #2 AND #3 AND #4
With the following limits:
- Map to preferred terminology (with spell check)
- Also search as free text
- Include sub-terms/derivatives (explosion search)
- Search publications: all years

Box 7. CENTRAL search string used for literature retrieval regarding the accuracy of HPV tests on self-samples.

#1: Cervix or cervico* or cervica*
#2: Cancer or carcinoma or neoplas* or dysplas* or CIN or CIN2 or CIN3 or SIL or SIL or HSIL or H-SIL or LSIL or L-SIL or "low grade" or low-grade or mild or equivocal or borderline.
#3: HPV or ‘‘human papillomavirus’’ or papillomavir* or viral or virus
#4: self-collection or "self collection" or self-sampling or ‘‘self-sampling’’ or self-collect* or self-sampl* or self
With the following limits:
- Cochrane reviews (reviews + protocols)
- Other reviews

Search for word variations

15.4. Studies included in the meta-analysis of the accuracy of HPV testing in triage of women with minor cytological abnormalities

15.4.1. Triage of ASC-US

A list of all studies included in the meta-analysis on triage of ASC-US is shown below. Studies in bold, are those that were not yet included in a previously published meta-analysis[69].

Test	No	1st author, Journal Year; Vol: Pages[Ref]
HC2		
	1	Manos, JAMA 1999; 281: 1605-1610 [373]
	2	Bergeron, Obstet Gynecol 2000; 95: 821-827 [374]
	3	Lytwyn, Can Med Ass J 2000; 19: 701-707 [375]
	4	Shlay, Obstet Gynecol 2000; 96: 410-416 [376]
	5	Morin, J Reprod Med 2001; 46: 799-805 [377]
	6	Rebello, BMJ 2001; 322: 893-894 [378]
	7	Solomon, J Natl Cancer Inst 2001; 93: 293-299 [379]
	8	Zielinski, J Pathol 2001; 195: 300-306 [380]
	9	Kulasingam, JAMA 2002; 288: 1749-1757 [234]
	10	Pretorius, J Reprod Med 2002; 47: 290-296 [381]
	11	Cuzick, Lancet 2003; 363: 1871-1876 [382]
	12	Guyot, BMC Infect Dis 2003; 3: 23-23 [383]
	13	Lonky, Obstet Gynecol 2003; 101: 481-489 [384]
	14	Ordi, Med Clin (Barc) 2003; 121: 441-445 [385]
	15	Wensveen, Acta Obstet Gynecol Scand 2003; 82: 883-889 [386]
	16	Andersson, Acta Obstet Gynecol Scand 2005; 84: 996-1000 [387]
	17	Chen , Taiw J Obstet Gynecol 2005; 44: 252-257 [388]
	18	Dalla Palma, Cytopathology 2005; 16: 22-26 [389]
	19	Davis-Devine, Am J Clin Pathol 2005; 124: 24-30 [390]
	20	Giovannelli, J Clin Virol 2005; 33: 281-286 [391]
	21	Nieh, Gynecol Oncol 2005; 97: 35-40 [392]
	22	Bergeron, Gyn Obstet Fert 2006; 34: 312-316 [393]
	23	Holladay, Cancer Cytopathol 2006; 108: 451-461 [394]
	24	Kelly, Cancer 2006; 108: 494-500 [395]
	25	Kiatpongsan, In J Gynecol Cancer 2006; 16: 262-26 5 [396]
	26	Monsonego, In J Gynecol Cancer 2006; 16: 591-598 [397]
	27	Cushieri, J Clin Virol 2007; 38: 14-18 [398]
	28	Ronco, Eur J Cancer 2007; 43: 476-480 [399]
	29	You, Aust NZ J Obstet Gynecol 2007; 47: 141-144 [400]
	30	De Francesco, J Virol Meth 2008; 147: 10-17 [401]
	31	Monsonego, Sex Transm Dis 2008; 35: 521-527 [402]
	32	Siddiqui, Arch Pathol Lab Med 2008; 132: 1648-1652 [403]
	33	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	34	Cattani, J Clin Microbiol 2009; 47: 3895-3901[405]
	35	Huang, J Clin Virol 2009; 45S: 19-23 [406]
	36	Lee, Int J Gynecol Cancer 2009; 19: 266-272 [407]
	37	Silverloo, Acta Obstet Gynaecol 2009; 88: 1006-1010 [408]
	38	Del Mistro, Gynecol Oncol 2010; 117: 77-88 [409]

Test	No	1st author, Journal Year; Vol: Pages[Ref]
	39	Denton, L Clin Pathol 2010; 134: 12-21 [152]
	40	**Halfon, Cancer Biomarkers 2010; 7: 133-139 [410]**
	41	**Alameda, Diagn Cytopathol 2011; 39: 110-11 4[411]**
	42	**Belinson, Am J Clin Pathol 2011; 135: 790-795 [267]**
	43	**CLEAR, http://www.accessdata.fda.gov/ 2011**
	44	**Clad, J Clin Microbiol 2011; 49: 1071-1076** [412]
	45	**Dufresne, J Clin Microbiol 2011; 49: 48-53 [413]**
	46	**Monsonego, Int J Cancer 2011; 129: 691-701 [328]**
	47	**Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]**
	48	**Schmidt, Cancer Cytopathol 2011; 119: 158-166 [315]**
	49	**Stoler, Am J Clin Pathol 2011; 135: 468-475 [415]**
	50	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
	51	**Alaghehbandan, Diagn Cytopathol 2013; 41: 767-775 [417]**
	52	**Oliveira, J Med Virol 2013; 85: 1235-1241[418]**
Amplicor		
	1	Monsonego, Gynecol Oncol 2005; 99: 160-168 [419]
	2	De Francesco, J Virol Meth 2008; 147: 10-17 [401]
	3	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	4	Wentzensen, CEBP2009; 18: 1341-1349 [420]
	5	Dufresne, J Clin Microbiol 2011; 49: 48-53 [413]
	6	**Jakobsson, Int J STD AIDS 2012; 23: 485-489 [421]**
Abott RT PCR		
	1	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	2	Huang, J Clin Virol 2009; 45S: 19-23 [406]
	3	Halfon, J Clin Virol 2010; 48: 246-250 [422]
	4	Wong, J Clin Virol 2011; 51: 136-138 [151]
	5	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
Linear Array		
	1	Castle, J Clin Microbiol 2008; 46: 109-117 [368]
	2	Fröberg, BJC 2008; 99: 563-568 [423]
	3	Monsonego, Sex Transm Dis 2008; 35: 521-527 [402]
	4	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	5	Lee, Int J Gynecol Cancer 2009; 19: 266-272 [407]
	6	Halfon, J Clin Virol 2010; 47: 38-42 [424]
	7	Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]
	8	Lapierre, J Clin Microbiol 2012; 50: 1240-1244 [425]
	9	**Dona, Gynecol Oncol 2012; 126: 198-202 [426]**
	10	Waldstrom, Cytopathol 2012; 23:389-395 [427]
	11	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**
PapilloCheck		
	1	Halfon, J Clin Virol 2010; 47: 38-42 [424]
Cervista		
	1	Wong, Cancer 2009; 115: 823-832 [429]
	2	Einstein, Gynecol Oncol 2010; 118: 116-122 [430]
	3	Bellinson, Am J Clin Pathol 2011; 135: 790-795 [267]
Cobas-4800		
	1	Stoler, Am J Clin Pathol 2011; 135: 468-475[415]
	2	Lapierre, J Clin Microbiol 2012; 50: 1240-1244 [425]

Test	No	1st author, Journal Year; Vol: Pages[Ref]
	3	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
Pretect-HPV Proofer		
	1	Molden, Int J Cancer 2005; 114: 973-976 [431]
	2	**Andersson, Int J Oncol 2006; 29: 705-711 [432]**
	3	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	4	Halfon, J Clin Virol 2010; 47: 177-181 [433]
	5	Sorbye , J Virol Meth 2010; 169: 219-222 [434]
	6	**Andersson, J Clin Microbiol 2011; 49: 3794-3799 [435]**
	7	Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]
	8	**Koliopoulos, Acta Obstet Gynecol Scand 2012; 91: 794-801 [436]**
	9	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
	10	**Alaghehbandan, Diagn Cytopathol 2013; 41: 767-775 [417]**
	11	**Oliveira, J Med Virol 2013; 85: 1235-1241[418]**
APTIMA		
	1	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	2	Dockter, J Clin Virol 2009; 45 (S1): S55-S61 [437]
	3	Wu, Int J Gynecol Cancer 2010; 20: 1411-1414 [438]
	4	Clad, J Clin Microbiol 2011; 49: 1071-1076 [412]
	5	CLEAR, http://www.accessdata.fda.gov/ 2011
	6	Monsonego, Int J Cancer 2011;129:691-701 [328]
	7	Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]
	8	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
	9	**Waldstrom, Cytopathol 2012; 23:389-395 [427]**
	10	**Stoler, Am J Clin Pathol 2011; 135: 468-475 [415]**
$p16^{INK4a}$		
	1	Nieh, Gynecol Oncol 2005; 97: 35-40 [439]
	2	**Andersson, Int J Oncol 2006; 29: 705-711 [432]**
	3	Holladay , Cancer 2006; 108: 451-461 [394]
	4	**Meyer, Cancer 2007; 111: 83-92 [440]**
	5	Monsonego, Acta Cytol 2007; 51: 755-766 [441]
	6	Wentzensen, Cancer 2007; 111: 58-66 [313]
	7	Schledermann, Diagn Cytopathol 2008; 36: 453-459 [442]
	8	Szarewski, CEBP2008; 17: 3033-3042 [404]
	9-11	Denton, Am J Clin Pathol 2010; 134: 12-21 (3 interpretations) [152]
	12	Guo, Diagn Cytopathol 2011; 39: 482-488 [443]
	13	Passamonti, Pathalogica 2010; 102: 6-11 [444]
	14	Samarawardana, Cancer Cytopathol 2010; 118: 146-156 [445]
	15	Sung, Diagn Cytopathol 2010; 38: 168-171 [446]
	16	Alameda, Diagn Cytopathol 2011; 39: 119-114 [411]
	17	Nasioutziki, Int J Gynecol Cancer 2011; 21: 79-85 [447]
	18	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
	19	**Loghavi, Diagn Cytopathol 2013; 41: 582-587 [448]**
P16/Ki-67		
	1	**Edgerton, Diagn Cytopathol 2011; 41: 35-40 [449]**
	2	**Schmidt, Cancer Cytopathol 2011; 119: 158-166 [315]**
	3	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**
	4	**Loghavi, Diagn Cytopathol 2013; 41: 582-587 [448]**

Test	No	1st author, Journal Year; Vol: Pages[Ref]	
HPV16 genotyping			Assay
	1	Froberg, Br J Cancer 2008; 99: 563-568 [423]	Linear array
	2	Monsonego, Int J STD AIDS 2008; 19: 385-392 [450]	Linear array
	3	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Linear array
	4	**Dona, Gynecol Oncol 2012; 126: 198-202 [426]**	Linear array
	5	Lapierre, J Clin Microbiol 2012; 50: 1240-1244 [425]	Linear array
	6	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**	Linear array
	7	**Gage, Cancer Epidemiol Biom Prev 2013; 22: 1095-1101 [451]**	Linear array
	8	Guo, Mod Pathol 2008; 21: 1037-1043 [452]	Easy Chip
	9	Huang, J Clin Virol 2009; 45: S19-S23 [406]	Easy Chip
	10	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Abbott
	11	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Abbott
	12	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	Abbott
	13	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Clinical Arrays
	14	Halfon, J Clin Virol 2010; 47: 177-181 [433]	PapilloCheck
	15	Wong, Cancer 2009; 115: 823-832 [429]	Cervista
	16	**Einstein, Gynecol Oncol 2010; 118: 116-122 [428]**	Cervista
	17	**Belinson, Am J Clin Pathol 2011; 135: 790-795 [267]**	MALDI-TOF
	18	**Depuydt, Cancer Epidemiol Biom Prev 2011; 20: 628-637 [453]**	E6/7 qPCR
	19	Stoler, Am J Clin Pathol 2011; 135: 468-475 [415]	COBAS-4800
	20	Lapierre, J Clin Microbiol 2012; 50: 1240-1244 [425]	COBAS-4800
	21	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	COBAS-4800
	22	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	BD Viper
	23	**Oliveira, J Med Virol 2013; 85: 1235-1241 [418]**	CLART
HPV16/18 genotyping			Assay
	1	Froberg, Br J Cancer 2008; 99: 563-568 [423]	Linear array
	2	Monsonego, Int J STD AIDS 2008; 19: 385-392 [450]	Linear array
	3	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Linear array
	4	**Dona, Gynecol Oncol 2012; 126: 198-202 [426]**	Linear array
	5	Lapierre, J Clin Microbiol 2012; 50: 1240-1244 [425]	Linear array
	6	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**	Linear array
	7	**Gage, Cancer Epidemiol Biom Prev 2013; 22: 1095-1101 [451]**	Linear array
	8	Guo, Mod Pathol 2008; 21: 1037-1043 [452]	Easy Chip
	9	Huang, J Clin Virol 2009; 45: S19-S23 [406]	Easy Chip
	10	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Abbott
	11	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Abbott
	12	Wong, J Clin Virol 2011; 51: 136-138 [151]	Abbott
	13	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	Abbott
	14	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Clinical Arrays
	15	Halfon, J Clin Virol 2010; 47: 177-181 [433]	PapilloCheck
	16	Wong, Cancer 2009; 115: 823-832 [429]	Cervista
	17	**Einstein, Gynecol Oncol 2010; 118: 116-122 [428]**	Cervista
	18	**Belinson, Am J Clin Pathol 2011; 135: 790-795 [267]**	MALDI-TOF
	19	**Depuydt, Cancer Epidemiol Biom Prev 2011; 20: 628-637 [453]**	E6/7 qPCR
	20	Stoler, Am J Clin Pathol 2011; 135: 468-475 [415]	COBAS-4800
	21	Lapierre, J Clin Microbiol 2012; 50: 1240-1244 [425]	COBAS-4800
	22	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	COBAS-4800
	23	**Spathis, Plos one 2012; 7: e49205 [454]**	CLART

Test	No	1st author, Journal Year; Vol: Pages[Ref]	
	24	**Oliveira, J Med Virol 2013; 85: 1235-41 [418]**	CLART
	25	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	BD Viper
ProExC			
	1	**Kelly, Cancer Cytopathol 2006; 108: 494-500 [395]**	
	2	**Siddiqui, Arch Pathol Lab Med 2008; 132: 1648-1652 [403]**	
	3	**Tambouret, Arch Pathol Lab Med 2008; 132: 918-925 [455]**	
	4	**Depuydt, Cancer Epidemiol Biom Prev 2011; 20: 628-637 [453]**	
	5	**Alaghehbandan, Diagn Cytopathol 2013; 41: 767-775 [417]**	

15.4.2. Triage of LSIL

A list of all studies included in the meta-analysis on triage of LSIL is shown below. Studies in bold, are those that were not yet included in a previously published meta-analysis[69].

Test	No	1st author, Journal Year; Vol: Pages[Ref]
HC2		
	1	Bergeron, Obstet Gynecol 2000; 95: 821-827 [374]
	2	Lytwyn, Can Med Ass J 2000; 19: 701-707 [375]
	3	**Lee, Arch Pathol Lab Med 2001; 125: 1453-1457 [456]**
	4	Rebello, BMJ 2001; 322: 893-894 [378]
	5	Zielinski, J Pathol 2001; 195: 300-306 [380]
	6	Kulasingam, JAMA 2002; 288: 1749-1757 [234]
	7	Pretorius, J Reprod Med 2002; 47: 290-296 [381]
	8	Sherman, J Natl Cancer Inst 2002; 94: 102-107 [127]
	9	Guyot, BMC Infect Dis 2003; 3: 23-2 3[383]
	10	Andersson, Acta Obstet Gynecol Scand 2005; 84: 996-1000 [387]
	11	Chen , Taiw J Obstet Gynecol 2005; 44: 252-257 [388]
	12	Holladay, Cancer Cytopathol 2006; 108: 451-461 [394]
	13	Monsonego, In J Gynecol Cancer 2006; 16: 591-598 [397]
	14	**Meyer, Cancer 2007; 111: 83-92 [440]**
	15	Ronco, Eur J Cancer 2007; 43: 476-480 [399]
	16	You, Aust NZ J Obstet Gynecol 2007; 47: 141-144 [400]
	17	De Francesco, J Virol Meth 2008; 147: 10-17 [401]
	18	Monsonego, Sex Transm Dis 2008; 35: 521-527 [402]
	19	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	20	Cattani, J Clin Microbiol 2009; 47: 3895-3901 [405]
	21	Huang, J Clin Virol 2009; 45S: 19-23 [406]
	22	Lee, Int J Gynecol Cancer 2009; 19: 266-272 [407]
	23	Castle , Obstet Gynecol 2010; 116: 76-84 [457]
	24	Denton, L Clin Pathol 2010; 134: 12-21 [152]
	25	**Halfon, Cancer Biomerkers 2010; 7: 133-139 [410]**
	26	Voss, Anal Quant Cytol Histol 2010; 32: 121-130 [458]
	27	**Wu, Int J Gynecol Cancer 2010; 20: 1411-1414 [438]**
	28	**Clad, J Clin Microbiol 2011; 49: 1071-1076 [412]**
	29	**Heider, Acta Cytol 2011; 55: 48-53 [459]**
	30	**Levi, Cancer Cytopathol 2011; 119: 228-234 [460]**
	31	**Monsonego, Int J Cancer 2011; 129: 691-701 [328]**

Test	No	1st author, Journal Year; Vol: Pages[Ref]
	32	**Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]**
	33	**Schmidt, Cancer Cytopathol 2011; 119: 158-166 [315]**
	34	**Tsoumpou, Gynecol Oncol 2011; 121: 49-53 [461]**
	35	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
	36	**Ziemke, Pathologe 2013; 34: 563-568 [462]**
	37	**Alaghehbandan, Diagn Cytopathol 2013; 41: 767-775 [417]**
	38	**Oliveira, J Med Virol 2013; 85: 1235-1241 [418]**
Amplicor		
	1	Monsonego, Gynecol Oncol 2005; 99: 160-168 [419]
	2	De Francesco, J Virol Meth 2008; 147: 10-1 7[401]
	3	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	4	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
Abott RT PCR		
	1	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	2	Huang, J Clin Virol 2009; 45S: 19-23 [406]
	3	Halfon, J Clin Virol 2010; 48: 246-250 [422]
	4	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
Linear Array		
	1	Fröberg, BJC 2008; 99: 563-568 [423]
	2	Monsonego, Sex Transm Dis 2008; 35: 521-527 [402]
	3	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	4	Lee, Int J Gynecol Cancer 2009; 19: 266-272 [407]
	5	Halfon, J Clin Virol 2010; 47: 38-42 [424]
	6	Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]
	7	**Dona, Gynecol Oncol 2012; 126: 198-202 [426]**
	8	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**
PapilloCheck		
	1	Jones, J Clin Virol 2009; 45: 100-104 [463]
	2	Halfon, J Clin Virol 2010; 47: 38-42 [424]
Cervista		
	1	Belinson, Am J Clin Pathol 2011; 135: 790-795 [267]
COBAS-4800		
	1	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
	2	**Cuzick, Int J Cancer 2013; 132: 959-966 [464]**
Pretect-HPV Proofer		
	1	Molden, Int J Cancer 2005; 114: 973-976 [431]
	2	**Andersson, Int J Oncol 2006; 29: 705-711 [432]**
	3	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	4	Halfon, J Clin Virol 2010; 47: 177-181 [433]
	5	Sorbye , J Virol Meth 2010; 169: 219-222 [434]
	6	Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]
	7	**Koliopoulos, Acta Obstet Gynecol Scand 2012; 91: 794-801 [436]**
	8	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**
	9	**Alaghehbandan, Diagn Cytopathol 2013; 41: 767-775 [417]**
	10	**Oliveira, J Med Virol 2013; 85: 1235-1241 [418]**
APTIMA		
	1	Szarewski,CEBP 2008; 17: 3033-3043 [404]
	2	Dockter, J Clin Virol 2009; 45 (S1): S55-S61 [437]

Test	No	1st author, Journal Year; Vol: Pages[Ref]	
	3	Wu, Int J Gynecol Cancer 2010; 20: 1411-1414 [438]	
	4	Clad, J Clin Microbiol 2011; 49: 1071-1076 [412]	
	5	Monsonego, Int J Cancer 2011; 129: 691-701 [328]	
	6	Ratnam, J Clin Microbiol 2011; 49: 557-564 [414]	
	7	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	
	8	**Waldstrom, Cancer Cytopathol 2013; 121: 136-145 [427]**	
ProExC			
	1	**Kelly, Cancer Cytopathol 2006; 108: 494-500 [395]**	
	2	**Tambouret, Arch Pathol Lab Med 2008; 132: 918-925 [455]**	
	3	**Depuydt, Cancer Epidemiol Biom Prev 2011; 20: 628-637 [453]**	
	4	**Alaghehbandan, Diagn Cytopathol 2013; 41: 767-775 [417]**	
$p16^{INK4a}$			
	1	**Andersson, Int J Oncol 2006; 29: 705-711 [432]**	
	2	Holladay , Cancer 2006; 108: 451-46 [394]	
	3	Meyer, Cancer 2007; 111: 83-92 [440]	
	4	Wentzensen, Cancer 2007; 111: 58-66 [313]	
	5	Schledermann, Diagn Cytopathol 2008; 36: 453-459 [442]	
	6	Szarewski, CEBP 2008; 17: 3033-3042 [404]	
	7-9	Denton, Am J Clin Pathol 2010; 134: 12-21 (3 interpretations)[152]	
	10	Passamonti, Pathalogica 2010; 102: 6-11 [444]	
	11	Samarawardana, Cancer Cytopathol 2010; 118: 146-156 [445]	
	12	Nasioutziki, Int J Gynecol Cancer 2011; 21: 79-85 [447]	
	13	Tsoumpou, Gynecol Oncol 2011; 121: 49-53 [461]	
	14	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	
	15	**Loghavi, Diagn Cytopathol 2013; 41: 582-587 [448]**	
p16/Ki-67			
	1	Schmidt, Cancer Cytopathol 2011; 119: 158-166 [315]	
	2	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**	
	3	**Ziemke, Pathologe 2013;34: 323-328 [462]**	
	4	**Loghavi, Diagn Cytopathol 2013; 41: 582-587 [448]**	
	5	**Waldstrom, Cancer Cytopathol 2013; 121: 136-145 [427]**	
HPV16 genotyping			Assay
	1	Froberg, Br J Cancer 2008; 99: 563-568 [423]	Linear array
	2	Monsonego, Int J STD AIDS 2008; 19: 385-392 [450]	Linear array
	3	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Linear array
	4	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Linear array
	5	**Dona, Gynecol Oncol 2012; 126: 198-202 [426]**	Linear array
	6	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**	Linear array
	7	**Gage, Cancer Epidemiol Biom Prev 2013; 22: 1095-1101 [451]**	Linear array
	8	Guo, Mod Pathol 2008; 21: 1037-1043 [452]	Easy Chip
	9	Huang, J Clin Virol 2009; 45: S19-S23 [406]	Easy Chip
	10	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Clinical Arrays
	11	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Abbott
	12	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Abbott
	13	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	Abbott
	14	Jones, J Clin Virol 2009; 45: 100-104 [463]	Papillocheck
	15	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Papillocheck
	16	Jones, J Clin Virol 2009; 45: 100-104 [463]	PCR

Test	No	1st author, Journal Year; Vol: Pages[Ref]	
	17	**Belinson, Am J Clin Pathol 2011; 135: 790-795 [267]**	MALDI-TOF
	18	**Depuydt, Cancer Epidemiol Biom Prev 2011; 20: 628-637 [453]**	E6/7 qPCR
	19	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	COBAS-4800
	20	**Cuzick, Int J Cancer 2013; 132: 959-966 [464]**	COBAS-4801
	21	**Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]**	BD-Viper
	22	**Oliveira, J Med Virol 2013; 85: 1235-1241 [418]**	CLART
HPV16/18 genotyping			Assay
	1	Froberg, Br J Cancer 2008; 99: 563-568 [423]	Linear array
	2	Monsonego, Int J STD AIDS 2008; 19: 385-392 [450]	Linear array
	3	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Linear array
	4	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Linear array
	5	**Dona, Gynecol Oncol 2012; 126: 198-202 [426]**	Linear array
	6	**Wentzensen, Clin Cancer Res 2012; 18: 4154-4162 [428]**	Linear array
	7	**Gage, Cancer Epidemiol Biom Prev 2013; 22: 1095-1101 [451]**	Linear array
	8	Guo, Mod Pathol 2008; 21: 1037-1043 [452]	EasyChip HPV Blot
	9	Huang, J Clin Virol 2009; 45: S19-S23 [406]	EasyChip HPV Blot
	10	Szarewski,CEBP 2008; 17: 3033-3043 [404]	Clinical Arrays
	11	Szarewski,CEBP 2008; 17: 3033-3044 [404]	Abbott
	12	Jones, J Clin Virol 2009; 45: 100-104 [463]	Papillocheck
	13	Halfon, J Clin Virol 2010; 47: 177-181 [433]	Papillocheck
	14	Jones, J Clin Virol 2009; 45: 100-104 [463]	PCR
	15	Belinson, Am J Clin Pathol 2011; 135: 790-795 [267]	MALDI-TOF
	16	Depuydt, Cancer Epidemiol Biom Prev 2011; 20: 628-637 [453]	E6/7 qPCR
	17	Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]	COBAS-4800
	18	Cuzick, Int J Cancer 2013; 132: 959-966 [464]	COBAS-4801
	19	Spathis, Plos one 2012; 7: e49205 [454]	CLART
	20	Oliveira, J Med Virol 2013; 85: 1235-1241 [418]	CLART
	21	Szarewski, J Clin Microbiol 2012; 50: 1867-1873 [416]	BD-Viper

16. Biography and bibliography

I obtained diplomas of Doctor in Medicine, Obstetrics and Surgery (Gent, 1984), Doctor in Tropical Medicine (Antwerp, 1985) and Masters in Tropical Biomedical Science (Antwerp, 1995). After a career in Africa with *Médecins Sans Frontières*, I got involved in cancer research in Europe. I coordinated the evaluation of new screening methods in the framework of the *European Network of Cervical Cancer Screening* and was editor of the second edition of the *European Guidelines for Cervical Cancer Screening* and associated Supplements on HPV screening. My main activity deals with systematic reviews, Cochrane reviews and meta-analyses regarding cervical cancer screening, diagnosis and treatment of screen detected cervical cancer precursors and HPV vaccination. Since the last years, we conducted also systematic reviews in the field of HPV-related head-and neck cancers. I am steering group member or coordinator of

several international networks involved in biobank-based cancer research that addresses etiological questions and possible application of biomarkers in screening, diagnosis and prognosis prediction and offer epidemiological support to health authorities and research organisations from about twenty countries regarding implementation of cervical cancer prevention. I am also specialised in trend analysis and age-cohort-period modelling.

Currently, I am the coordinator of the Unit of Cancer Epidemiology, which is part of the Belgian Cancer Centre at the Scientific Institute of Public Health in Brussels.

I am author or co-author of more than 200 scientific papers, published in peer-reviewed journals and a few dozens of book chapters. In Table 32, the twenty most cited papers are listed. My h-index in www.scopus.com is 51, which means that I (co-) authored 46 papers that were cited at least 51 times. I presented approximately 400 oral communications at international conferences, most of them as invited speaker. I rank at the third place of most influential leaders on cervical cancer worldwide on http://www.expertscape.com/leaders/cervical-cancer.

Table 32. Twenty most cited scientific papers of M. Arbyn (www.scopus.com 28/10/2017).

Rank	1st author	Year	Title	Journal	Citations
1	Kyrgiou M, et al	2006	Obstetric outcomes after conservative treatment for intraepithelial or early invasive cervical lesions: Systematic review and meta-analysis	Lancet	703
2	Arbyn M, et al	2011	Worldwide burden of cervical cancer in 2008	Ann Oncol	526
3	Cuzick J, et al	2008	Overview of Human Papillomavirus-Based and Other Novel Options for Cervical Cancer Screening in Developed and Developing Countries	Vaccine	405
4	Arbyn M, et al	2006	Chapter 9: Clinical applications of HPV testing: A summary of meta-analyses	Vaccine	369
5	Ronco G, et al	2014	Efficacy of HPV-based screening for prevention of invasive cervical cancer: Follow-up of four European randomised controlled trials	Lancet	361
6	Arbyn M, et al	2008	Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis.	BMJ	342
7	Arbyn M, et al	2004	Virologic versus cytologic triage of women with equivocal pap smears: A meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia	J Natl Cancer Inst	309
8	Meijer C, et al	2009	Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older	Int J Cancer	289
9	Arbyn M, et al	2008	Liquid compared with conventional cervical cytology: A systematic review and meta-analysis	Obstet Gynecol	279
10	Arbyn M, et al	2012	Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer	Vaccine	266
11	Arbyn M, et al	2010	European guidelines for quality assurance in cervical cancer screening. Second edition-summary document	Annf Oncol	202
12	Bray F, et al	2005	Trends in cervical squamous cell carcinoma incidence in 13 European countries: Changing risk and the effects of screening	Cancer Epidem Biom Prev	198
13	Tsoumpou I, et al	2009	p16^{INK4a} immunostaining in cytological and histological specimens from the uterine cervix: A systematic review and meta-analysis	Cancer Treat Rev	194
14	Arbyn M, et al	2005	Clinical utility of HPV-DNA detection: Triage of minor cervical lesions, follow-up of women treated for high-grade CIN: An update of pooled evidence	Gynecol Oncol	168
15	Koliopoulos G, et al	2007	Diagnostic accuracy of human papillomavirus testing in primary cervical screening: A systematic review and meta-analysis of non-randomized studies	Gynecol Oncol	167
16	Arbyn M, et al	2014	Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: A meta-analysis	Lancet Oncol	149
17	Ndiaye C, et al	2014	HPV DNA, E6/E7 mRNA, and p16^{INK4a} detection in head and neck cancers: A systematic review and meta-analysis	Lancet Oncol	145
18	Arbyn M, et al	2008	Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India	Int J Cancer	141
19	Anttila A, et al	2004	Cervical cancer screening programmes and policies in 18 European countries	Brit J Cancer	141
20	Paraskevaidis E, et al	2004	The role of HPV DNA testing in the follow-up period after treatment for CIN: A systematic review of the literature	Cancer Treat Rev	139
21-236					4,866
				Total	**10,359**

In grey: papers including systematic reviews and meta-analyses.

17. Attached publications

1. Arbyn M, Buntinx F, Van Ranst M, Paraskevaidis E, Martin-Hirsch P, Dillner J. Virologic versus cytologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. J Natl Cancer Inst 2004; 96: 280-93.

2. Arbyn M, Ronco G, Anttila A, Meijer CJLM, Poljak M, Ogilvie G, Koliopoulos G, Naucler P, Sankaranarayanan R, Peto J. Evidence regarding HPV testing in secondary prevention of cervical cancer. Vaccine 2012; 30 Suppl 5: F88-F99.

3. Arbyn M, Roelens J, Cuschieri K, Cuzick J, Szarewski A, Ratnam S, Reuschenbach M, Belinson J, Belinson JL, Monsonego J. The APTIMA HPV assay versus the Hybrid Capture II test in triage of women with ASC-US or LSIL cervical cytology: a meta-analysis of the diagnostic accuracy. Int J Cancer 2013; 132: 101-8.

4. Arbyn M, Verdoodt F, Snijders PJF, Verhoef VM, Suonio E, Dillner L, Minozzi S, Bellisario C, Banzi R, Zhao FH, Hillemans P, Anttila A. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. Lancet Oncol 2014; 15: 172-83.

5. Arbyn M, Vanden Broeck D, Benoy I, Bogers J, Depuydt C, Praet M, De Sutter P, Hoorens A, Hauben E, Poppe W, Van Ranst M, Delvenne P, Gofflot S, Petain M, Engelen F, Vanneste A, Op De Beeck L, Van Damme P, Temmerman M, Weyers S. Surveillance of effects of HPV vaccination in Belgium. Cancer Epidemiol 2016; 41: 152-8.

6. Arbyn M, Xu L, Verdoodt F, Cuzick J, Szarewski A, Belinson J, Wentzensen N, Gage JC, Khan MJ. Genotyping for human papillomavirus types 16 and 18 in women with minor cervical lesions: a systematic r eview and meta-analysis. Ann Intern Med 2017; 166: 118-127.

Virologic Versus Cytologic Triage of Women With Equivocal Pap Smears: A Meta-analysis of the Accuracy To Detect High-Grade Intraepithelial Neoplasia

Marc Arbyn, Frank Buntinx, Marc Van Ranst, Evangelos Paraskevaidis, Pierre Martin-Hirsch, Joakim Dillner

Background: **The appropriate management of women with minor cytologic lesions in their cervix is unclear. We performed a meta-analysis to assess the accuracy of human papillomavirus (HPV) DNA testing as an alternative to repeat cytology in women who had equivocal results on a previous Pap smear.** ***Methods:*** **Data were extracted from articles published between 1992 and 2002 that contained results of virologic and cytologic testing followed by colposcopically directed biopsy in women with an index smear showing atypical cells of undetermined significance (ASCUS). Fifteen studies were identified in which HPV triage and the histologic outcome (presence or absence of a cervical intraepithelial neoplasia of grade II or worse [CIN2+]) was documented. Nine, seven, and two studies also documented the accuracy of repeat cytology when the cutoff for abnormal cytology was set at a threshold of ASCUS or worse, low-grade squamous intraepithelial lesion (LSIL) or worse, or high-grade squamous intraepithelial lesion (HSIL) or worse, respectively. Random-effects models were used for pooling of accuracy parameters in case of interstudy heterogeneity. Differences in accuracy were assessed by pooling the ratio of the sensitivity (or specificity) of HPV testing to that of repeat cytology.** ***Results:*** **The sensitivity and specificity were 84.4% (95% confidence interval [CI] = 77.6% to 91.1%) and 72.9% (95% CI = 62.5% to 83.3%), respectively, for HPV testing overall and 94.8% (95% CI = 92.7% to 96.9%) and 67.3% (95% CI = 58.2% to 76.4%), respectively, for HPV testing in the eight studies that used the Hybrid Capture II assay. Sensitivity and specificity of repeat cytology at a threshold for abnormal cytology of ASCUS or worse was 81.8% (95% CI = 73.5% to 84.3%) and 57.6% (95% CI = 49.5% to 65.7%), respectively. Repeat cytology that used higher cytologic thresholds yielded substantially lower sensitivity but higher specificity than triage with the Hybrid Capture II assay. The ratio of the sensitivity of the Hybrid Capture II assay to that of repeat cytology at a threshold of ASCUS or worse pooled from the four studies that used both triage tests was 1.16 (95% CI = 1.04 to 1.29). The specificity ratio was not statistically different from unity.** ***Conclusion:*** **The published literature indicates that the Hybrid Capture II assay has improved accuracy (higher sensitivity, similar specificity) than the repeat Pap smear using the threshold of ASCUS for an outcome of CIN2+ among women with equivocal cytologic results. The sensitivity of triage at higher cytologic cutoffs is poor. [J Natl Cancer Inst 2004;96:280–93]**

The main purpose of cytologic screening by Papanicolaou (Pap) smears is to identify women with cervical lesions that confer an increased risk of cervical cancer *(1).* Such women require further follow-up and/or treatment to prevent progression to invasive disease. Consensus exists that women with high-grade cytologic lesions should be referred immediately for further exploration *(2).* However, the management of women with minor cytologic lesions remains controversial *(3).* Follow-up recommendations for women with atypical squamous cells of undetermined significance (ASCUS) and low-grade squamous intraepithelial lesions (LSILs) vary from conservative repeat cytology *(4–7)* to immediate referral for colposcopy and biopsy *(8–11).* Referral to the colposcopy clinic and the subsequent histologic examination result in substantial costs for the health care system *(12)* and often create feelings of anxiety and discomfort for women *(13).*

The natural history of minor cytologic lesions is difficult to predict on the basis of cytomorphologic grounds. These lesions often regress spontaneously and do not require treatment *(14–16).* Referring all women with minor cytologic lesions for further gynecologic exploration would, therefore, mean an increase in overdiagnosis and overtreatment *(12,17).* Lack of availability of colposcopic services at affordable prices often makes such an approach unrealistic. Nevertheless, although most women with an ASCUS smear result do not have clinically significant disease, a substantial proportion of them do have histopathologically confirmed high-grade cervical intraepithelial neoplasia or worse (CIN2+) *(18–20).* Indeed, from a population of screened U.S. women, it was estimated that one-third of CIN2+ lesions were discovered on follow-up of a previous smear with ASCUS *(20).*

Given the evidence concerning the etiologic role of oncogenic human papillomavirus (HPV) infections in the development of cervical cancer and CIN *(21–25),* HPV testing has been proposed as a triage method to identify women at increased risk of cervical cancer and thus requiring referral for colposcopic exploration *(18,19).* However, results concerning the utility of HPV triage for women with equivocal cytology are inconsistent *(26–28).* We therefore used meta-analytic tools to extract from

Affiliations of authors: European Network for Cervical Cancer Screening and Unit for Evaluation of Cancer Screening Programmes, Scientific Institute of Public Health, Brussels, Belgium (MA); Centre for Evidence Based Medicine, Belgian Branch of the Cochrane Collaboration, Department of General Practice, University of Leuven, Leuven, Belgium, and University of Maastricht, Maastricht, The Netherlands (FB); Department of Clinical and Epidemiological Virology, University of Leuven (MVR); Department of Obstetrics and Gynaecology, University Hospital of Ioannina, Ioannina, Greece (EP); Central Lancashire Teaching Hospitals, Preston, UK (PMH); Department of Medical Microbiology, Lund University, Malmö University Hospital, Malmö, Sweden (JD).

Correspondence to: Marc Arbyn, MD, MSc, Scientific Institute of Public Health, J. Wytsmanstreet 14, B1050 Brussels, Belgium (e-mail: marc.arbyn@iph.fgov.be).

See "Notes" following "References."

DOI: 10.1093/jnci/djh037

Journal of the National Cancer Institute, Vol. 96, No. 4, © Oxford University Press 2004, all rights reserved.

the literature all available data concerning diagnostic accuracy parameters, studied the variation in a systematic way, and obtained overall synthetic measures.

METHODS

Research Question

In this meta-analysis, we addressed the following questions: 1) What is the accuracy (i.e., sensitivity, specificity, predictive values, and likelihood ratios) of HPV DNA testing to detect histologically confirmed CIN2+ disease in women with an index smear showing ASCUS? 2) In studies in which the Pap smear was repeated, what is the accuracy of repeat cytology at thresholds of ASCUS, LSIL, and HSIL (high-grade squamous intraepithelial lesion) to detect CIN2+? and 3) What are the differences in accuracy between both triage tests?

Retrieval Strategy

A systematic literature search identified articles published between 1992 and 2002 that contained quantitative data allowing assessment of one or more of the research questions. Articles were retrieved from the electronic bibliographic databases (MEDLINE, EMBASE, Cochrane Library) using the following search terms: (((cervix OR cervical) AND (cancer OR carcinoma OR neoplas* OR dysplas* OR CIN OR SIL)) OR (cervix neoplasm)) AND (HPV OR human papillomavirus) AND (triage OR management). The search was completed manually by searching the reference lists of relevant articles and by screening the tables of contents (for 1992–2002) of the following journals: *American Journal of Obstetrics and Gynecology*, *European Journal of Gynecological Oncology*, *Journal of Gynecological Oncology*, *Journal of Lower Genital Tract Disease*, *Journal of Reproductive Medicine*, and *Obstetrics and Gynecology*. References were selected if they fulfilled three inclusion criteria: 1) women in the study presented with an index Pap smear of the uterine cervix with atypical squamous/glandular cells of unspecified significance (ASCUS/AGUS); 2) an HPV DNA detection test was performed; and 3) women were subsequently subjected to colposcopy and colposcopy-directed biopsies, with or without endocervical curettage, for histologic verification. A fourth but non-obligatory criterion was the repetition of the Pap test.

For several studies, we considered limited information. From Ferris et al. *(29,30)*, we selected the HPV DNA data from the Hybrid Capture I assay because these data were contrasted with repeat cytology; from Morin et al. *(31)*, HPV data from the Hybrid Capture II assay were used; and from the ASCUS-LSIL Triage Study (ALTS) *(32,33)*, we used results from two of the three experimental arms: women randomly assigned to immediate colposcopic verification and women randomly assigned to the HPV DNA testing arm, in which colposcopy was restricted to women showing presence of high-risk HPV DNA or HSIL on the repeat smear *(32)*.

Thresholds for Triage Tests

We considered a single threshold for HPV positivity: the presence of HPV DNA at a level greater than the cutoff for a positive test as stated by the test manufacturer. The variation in accuracy according to other thresholds of increasing viral load was studied by a summary receiver operating characteristic curve analysis and is reported elsewhere *(34)*. We retrieved information on the presence of oncogenic and high-risk HPV types only. We considered three threshold levels for abnormal cytology on the repeat Pap test: ASCUS or worse, LSIL or worse, and HSIL or worse. The 1991 version of the Bethesda Reporting System *(35)* was used for cytologic classification.

Outcome

The histologic result was used as the gold standard. We assumed that histologic examination of material obtained by colposcopy-directed biopsy, loop excision, or endocervical curettage provided complete ascertainment of the considered disease status. Throughout our systematic review, we used the CIN nomenclature to describe histologic outcomes *(36)*. We considered only the outcome of CIN2+.

Covariate Information

The following study properties were summarized in comprehensive tables: characteristics of the study population (place, inclusion and exclusion criteria, study size, and age distribution), properties of the HPV DNA testing systems (type of DNA detection method, collection device, transport medium), repeat cytology (conventional or liquid-based preparation, cytologic threshold, collection device), procedures for gold standard verification, and the blinding of interpreters for other test results.

Definition of Accuracy Measures and Statistical Analysis

The numbers of true-positives, false-negatives, false-positives, and true-negatives defined at the considered thresholds were extracted from each study, and the following accuracy parameters were calculated: sensitivity (true-positives/[true-positives plus false-negatives]); specificity (true-negatives/[true-negatives plus false-positives]); positive predictive value (true-positives/[true-positives plus false-positives]); negative predictive value (true-negatives/[true-negatives plus false-negatives]); test-positivity rate ([true-positives plus false-positives]/total number of patients), and prevalence of disease defined as presence of CIN2+ ([true-positives plus false-negatives]/total number of patients). The positive likelihood ratio (PLR = sensitivity/[1 – specificity]) and the negative likelihood ratio (NLR = [1 – sensitivity]/specificity) express the likelihood of the presence of CIN2+ versus the absence of CIN2+ when tests are positive or negative, respectively. The positive likelihood ratio should be greater than unity and as large as possible, whereas the negative likelihood ratio should be less than unity and tend toward zero. To assess differences in accuracy, we determined the ratio of the sensitivity (or specificity) of HPV testing to that of repeat cytology for those studies where both test systems were evaluated on the same women and then pooled the individual ratios. The variation in accuracy measures in the individual studies and in the pooled measures were displayed graphically using forest plots *(37–39)*. Random-effects models were used for pooling accuracy parameters in cases of statistically significant interstudy heterogeneity (i.e., when $P<.10$ for Cochran's Q test) *(40,41)*. In the absence of statistically significant heterogeneity, fixed models were used, with weighting of each individual study parameter according to the reciprocal of its variance *(39)*. Meta-analyses were performed using the Stata statistical package (version 7.0; Stata Corp., College Station, TX) *(42)*. Subgroup meta-analysis was used to assess the influ-

ence of study characteristics on the outcome. Age-stratified data on accuracy to detect CIN2+ were published in only one study *(43)*. We also obtained age-stratified data from the ALTS directly from Dr. M. Schiffman (National Cancer Institute, Bethesda, MD) *(33)*.

Publication Bias

Publication bias generally arises when smaller studies have a higher chance of being published if their results are positive. We assessed publication bias by using the asymmetry regression test and asymmetry plots, in which the normal deviate of the accuracy measure is plotted against its precision, which is related to the study size *(44,45)*.

RESULTS

Study Characteristics

Initially, we identified 29 potential articles for inclusion *(18,19,29–33,43,46–66)*. However, we excluded 12 articles concerning triage of patients with ASCUS or LSIL, because the group with an ASCUS index smear could not be separated out *(55–66)*. In total, 17 articles reporting results of 15 studies met the inclusion criteria *(18,19,29–33,43,46–54)*. Among the 15 studies, data allowing computation of the accuracy of repeat cytology at the thresholds ASCUS or worse, LSIL or worse, or HSIL or worse were obtained from nine, seven, and two studies, respectively. The characteristics of the included studies are summarized in Table 1 and Table 2.

Study Size

In total, 5454 women in the 15 included studies had ASCUS triage with HPV DNA testing. Nine studies were small, each contributing fewer than 200 women; four studies were of intermediate size, each contributing between 200 and 500 women; and two studies were large, each contributing more than 500 women. One of the large studies, the ALTS, contributed more than 2300 women *(32,33)*.

Clinical Setting and Population Characteristics

For each study, patients were recruited from colposcopy clinics or from gynecologic services to which women had been referred because of a cytologic result of ASCUS. In four studies *(48,51,53,54)*, the referred women had had repeated atypical cytology. Six studies *(18,29,32,49,52,54)* excluded women with a history of CIN, cervical surgery, or biopsy. One study *(19)* presented separate results for women with and without previous CIN. One study *(31)* included only women with ASCUS occurring after two sequential normal smears.

Blinding and Quality Review of Histologic Outcome

In six studies *(18,29,43,47,49,52)*, it was explicitly stated that the histologic interpretation was blinded to the triage test results. In three studies *(18,32,49)*, the histopathologic diagnosis was subjected to review by expert histologists.

Triage Tests

Several different HPV detection methods were used in the included studies. Two studies *(46,47)* used ViraPap and ViraType. These are two older commercial DNA dot blot test kits that use radioactive HPV RNA probes. The ViraPap contains a cocktail of seven probes that detect HPV types 6, 11, 16, 18, 31, 33, and 35. ViraType contains two distinct high-risk HPV cocktails that detect the high-risk HPV types 16, 18, 31, 33, and 35.

Hybrid Capture techniques were the most frequently used HPV DNA triage methods. The first generation Hybrid Capture I assay or Hybrid Capture Tube, which detects nine high-risk HPV types (i.e., 16, 18, 31, 33, 35, 45, 51, 52, and 56), was evaluated in five studies *(18,19,29,48,51)*. The Hybrid Capture II assay, which contains a cocktail of probes that detect the same nine high-risk HPV types and HPV types 39, 58, 59, and 68 *(67)*, was used in eight studies *(31,32,43,49,50,52–54)*. In one study *(49)*, a prototype Hybrid Capture II assay was used that did not contain probes for HPV types 59 and 68. For all Hybrid Capture techniques, the hybridization yields a chemiluminescent signal that is compared with positive controls containing a known amount of HPV16 DNA. The detection limit is 10 pg of HPV DNA/mL in the Hybrid Capture I assay and 1 pg/mL in the Hybrid Capture II assay.

Cytology and Hybrid Capture II accuracy data from the same subjects were available in four studies *(31,32,49,50)*. The conventional Pap smear was used as the cytologic triage method in most studies, although the ThinPREP liquid-based technique was used in two studies *(32,49)*.

Triage by HPV Testing

The sensitivity to detect CIN2+ varied from extremely low, 26.7% with the ViraPap test *(47)*, to 95.9% *(32)* and 100.0% *(52)* with the Hybrid Capture II assay (Table 3). The negative predictive values ranged from 88.7% *(47)* to 100.0% *(52)*. The specificity ranged from 48.4% *(32)* to 97.1% *(51)*. The positive predictive values ranged from 7.8% *(19)* to 90.6% *(51)*. The prevalence of HPV positivity varied between 16.4% *(46,48)* and 56.8% *(32)*.

Triage by Repeat Cytology at Cutoff ASCUS or Worse

The numbers of true- and false-positive and true- and false-negative results for repeat cytology, defined at the cutoff ASCUS or worse, and derived parameters are shown in Table 4. Sensitivity ranged from 60.0% *(46)* to 85.0% *(32)*, and specificity ranged from 44.7% *(29,32)* to 71.7% *(50)*. Positive predictive values ranged from 3.8% *(46)* to 22.2% *(47,50)*, and negative predictive values ranged from 93.4% *(47)* to 97.8% *(46)*. The test positivity rate ranged from 32.4% *(50)* to 58.8% *(32)*.

Triage by Repeat Cytology at Cutoff LSIL or Worse

The numbers of true- and false-positive and true- and false-negative results for repeat cytology, defined at the cutoff LSIL or worse, and derived parameters are shown in Table 4. Sensitivity ranged from 20.0% *(29)* to 59.2% *(32)*, and specificity ranged from 77.9% *(32)* to 96.4% *(46)*. Because of higher specificity in triaging at cutoff LSIL, the positive predictive value increased and ranged from 8.0% *(29)* to 32.5% *(49)*. The negative predictive value ranged from 93.6% *(32)* to 98.2% *(46)*. Test positivity was considerably lower than triage at cutoff ASCUS and varied between 4.7% *(46)* and 26.4% *(32)*.

Triage by Repeat Cytology at Cutoff HSIL or Worse

Both studies *(32,49)* in which the accuracy of the repeat cytology was defined at the cutoff HSIL or worse had very low

Table 1. Characteristics of the study populations recruited for included triage studies*

Author, reference	Country	Study location	Study population and inclusion criteria	Exclusion criteria	Study size	Age distribution, y
Goff et al., 1993 *(46)*	United States	Colposcopy clinic in Boston, MA	Selected from 360 consecutive patients consulting at a colposcopy clinic, with an ASCUS index smear	No information provided	171	Mean = 35; no further details
Slawson et al., 1994 *(47)*	United States	General practitioners in Harrisburg, PA	Selected from a GP screening population, with an ASCUS index smear	Incomplete data, refusal of colposcopy, women claiming immediate colposcopy	121	Mean = 25; range = 15–45
Cox et al., 1995 *(18)*	United States	University of California clinic in Santa Barbara	Young women referred because of ASCUS	Preference for follow-up with cytology, refusal of colposcopy, history of treatment for CIN	217	Mean = 21; >90% in 18–22 range
Wright et al., 1995 *(19)*	United States, Canada	Colposcopy clinics in New York and Quebec	Women referred to a colposcopy clinic (19% were previously treated for CIN), with index smear showing ASCUS/AGUS	Results are presented for all women (n = 181) and for the subgroup of women without history of treatment (n = 136)	181	Median = 36
Fait et al., 1998 *(48)*	Israel	Colposcopy clinic in Tel Aviv	Women with repeated ASCUS and normal previous colposcopy	No information provided	67	Mean = 28; range = 17–49
Ferris et al., 1998 *(29,30)*	United States	Six colposcopy clinics	Women referred to colposcopy clinic with a previous ASCUS <4 mo earlier	Pregnancy, HSIL, or cervical cancer <12 mo earlier, immunosuppression, prior colposcopy or CIN in last 12 mo; cervical evaluation impossible	169	Mean = 27; SD = 2.4; range = >18
Manos et al., 1999 *(49)*	United States	12 gynecology clinics belonging to the Kaiser Permanente Medical Care Programme (Northern California)	Screened women recalled for a smear showing ASCUS, consenting with participation for triage, and gold standard results were available	Pregnancy, treatment for cervical neoplasia <6 mo earlier, recalling impossible	973	Mean = 40; range = 14–92
Bergeron et al., 2000 *(50)*	France	41 private gynecologists in Paris	Screened women with index smear showing ASCUS	Biopsy containing only endocervical cylindrical cells	111	Mean = 35; SD = 10; range = 15–75
Fait et al., 2000 *(51)*	Israel	Colposcopy clinic in Tel Aviv	Women with two consecutive smears showing ASCUS	Patients with colposcopically visible HPV-related disease in the anogenital area	226	Mean = 28; range = 17–48
Lin et al., 2000 *(52)*	Taiwan	Colposcopy clinic in Chang Gung	Women with ASCUS <3 mo earlier	Menopausal bleeding <6 mo earlier, previous diagnosis or treatment of CIN or cancer, immunosuppression, gynecologic evaluation impossible	74	Median = 62; range = 50–78
Shlay et al., 2000 *(43)*	United States	Women's clinic in Denver, CO	Indigent women referred for ASCUS or AGUS	Menstruation, pregnancy, refusal of consent	195	≤20 (6.2%), 20–29 (32.8%), 30–39 (29.7%), ≥40 (31.3%)
Morin et al., 2001 *(31)*	Canada	Gynecology services, University of Quebec	Women with diagnosis of ASCUS after two consecutive normal smears; informed consent	Pregnancy, previous biopsy or treatment of the cervix	360	Range = 18–50
Rebello et al., 2001 *(53)*	United Kingdom	Colposcopy clinic in Edinburgh	Women with persistent borderline smears	No information provided	75	Median = 30; range = 17–61
Solomon, 2001 *(32)*; Sherman et al., 2002 *(33)*	United States	Gynecology clinics in Birmingham, AL; Oklahoma City, OK; Pittsburgh, PA; Seattle, WA	3600 women enrolled on average 2 mo after an index smear showing ASCUS enrolled in ALTS in the immediate colposcopy arm or in the HPV arm	Pregnancy, ablative surgery, or excisional therapy on the cervix	2336	Mean = 29
Zielinski et al., 2001 *(54)*	The Netherlands	General practitioners and outpatient clinic in Walcheren	278 women consulting at general practitioners or outpatient clinic with two previous smears showing borderline or mild dysplasia; 213 had a last smear showing borderline dysplasia	Known history of cervical pathology, glandular lesion, lost to follow-up	213	Mean = 41; range = 20–76

*ASCUS = atypical squamous cells of undetermined significance; AGUS = atypical glandular cells of undetermined significance; CIN = cervical intraepithelial neoplasia; HSIL = high-grade squamous intraepithelial lesion; HPV = human papillomavirus; ALTS = ASCUS-LSIL Triage Study.

Table 2. Characteristics of test procedures applied in retrieved triage studies*

Author	HPV triage test	Collection device for HPV	Transport medium for HPV	Cytologic triage test	Cutoff for cytologic triage	Collection device for cytology	Gold standard	Blinding of testing
Goff et al., 1993 *(46)*	ViraType	No information provided	No information provided	Pap smear	≥ASCUS ≥LSIL	No information provided	Colposycopy, ECC, colposcopically oriented biopsies when indicated	No information provided
Slawson et al., 1994 *(47)*	ViraPap	Dacron swab	ViraPap	Pap smear	≥ASCUS	No information provided	Colposcopy and directed biopsies. ECC for all cases. All histological results were reviewed	Histological review was blinded
Cox et al., 1995 *(18)*	HC1	Cone brush	ViraPap	Pap smear	≥ASCUS ≥LSIL	Ayre spatula and Cytobrush	Colposcopy and directed biopsies from abnormal zones. ECC if index smear showed glandular atypia or if atypical transformation zone	Histological review was blinded
Wright et al., 1995 *(19)*	HCT	Dacron swab	ViraPap	Pap smear	≥ASCUS ≥LSIL	Ayre spatula and Cytobrush	Colposcopy and biopsy of visible abnormalities ECC for most cases, also when no visible lesion was observed	No information provided
Fait et al., 1998 *(48)*	HC1	Cotton tip applicator	STM	None	Not applicable	Not applicable	Colposcopy and cone biopsies	No information provided
Ferris et al., 1998 *(29,30)*	HCT	Dacron swab	ViraPap	Pap smear	≥ASCUS ≥LSIL	Ayre spatula and Cytobrush or Ancellon sampler	Biopsies from lesions under colposcopic control. ECC when indicated	Histology was blinded from HPV
Manos et al., 1999 *(49)*	HC2	Papette brush	ThinPrep	Pap smear (and LBC)	≥ASCUS ≥LSIL ≥HSIL	Papette brush; Cytobrush if cervical stenosis	Colposcopy with biopsy and/or ECC for all cases. ECC for cases in which no lesion requiring biopsy was seen. Double (and if discordance, triple) reading of histology	Colposcopy/histology was blinded from triage test results
Bergeron et al., 2000 *(50)*	HC2, PCR, SB	Cone brush	STM	Pap smear	≥ASCUS	Ayre spatula and Cytobrush	Colposcopy and biopsies taken from the transformation zone	Pathologist was masked to the other test results
Fait et al., 2000 *(51)*	HCl	Dacron swab	STM	None	Not applicable	Not applicable	Punch biopsy from the most severe lesion, diagnostic cone biopsy if colposcopy was normal and adequate	Colposcopy not blinded; histology blinded
Lin et al., 2000 *(52)*	HC2	Cone brush	STM	None	Not applicable	Not applicable	Punch biopsies from most affected area, random biopsies and/or ECC when no visible lesion was identified. Treatment of CIN2+ with cervical cone or LEEP	Histology blinded to HPV test results
Shlay et al., 2000 *(43)*	HC2	Cone brush	STM	None	Not applicable	Not applicable	Biopsy from colposcopically abnormal areas. ECC and endocervical brush sample from all subjects	Histology blinded to clinic and HPV results
Morin et al., 2001 *(31)*	HC1/2 PCR	Dacron swab	STM	Pap smear	≥ASCUS ≥LSIL	No information provided	Colposcopy and directed biopsies or ECC for all lesions of the cervix	No information provided
Rebello et al., 2001 *(53)*	HC2	Cone brush	STM	None	Not applicable	Not applicable	Large loop excision on all patients	No information provided
Solomon et al., 2001 *(32)*; Sherman et al., 2002 *(33)*	HC2	Papette brush	ThinPrep	LBC	≥ASCUS ≥LSIL ≥HSIL	Papette brush	Biopsy from colposcopically suspected CIN, ECC if indicated (transformation zone or extent of lesion not completely visible). CIN2+ lesions treated with LEEP. Quality review of histology.	Histology not blinded from cytology. Colposcopy repeated when histology was less severe than cytology. No statement on independence of histology and/or virology. In arm 2, almost all referred cases were HPV-positive
Zielinski et al., 2001 *(54)*	HC2	No information provided	No information provided	None	Not applicable	Not applicable	Colposcopically directed biopsies when a lesion was visible	No information provided

*HC = Hybrid Capture; HCT = Hybrid Capture Tube; SB = Southern blot; PCR = polymerase chain reaction; ECC = endocervical curettage; LEEP = loop electrosurgical excision procedure; STM = specimen transport medium; LBC = liquid-based cytology; ASCUS = atypical squamous cells of undetermined significance; LSIL = low-grade squamous intraepithelial lesion; HSIL = high-grade squamous intraepithelial lesion; CIN = cervical intraepithelial neoplasia; HPV = human papillomavirus; ≥ = cutoff level or worse.

Table 3. Triage of atypical squamous cells of undetermined significance (ASCUS) by human papillomavirus (HPV) DNA testing for the detection of histologically confirmed cervical intraepithelial neoplasia grade II or worse (CIN2+): number of true- and false-positives and true- and false-negatives, accuracy parameters, test positivity rate, and prevalence of disease derived from different published studies*

Study	Type	True-positive	False-negative	False-positive	True-negative	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Test positivity rate	Prevalence of disease
Goff et al., 1993 *(46)*	VT	3	2	25	141	0.600	0.849	0.107	0.986	0.164	0.029
Slawson et al., 1994 *(47)*	VP	4	11	20	86	0.267	0.811	0.167	0.887	0.198	0.124
Cox et al., 1995 *(18)*	HC1	14	1	67	135	0.933	0.668	0.173	0.993	0.373	0.069
Wright et al., 1995 *(19)*	HC1	6	5	71	99	0.545	0.582	0.078	0.952	0.425	0.061
Fait et al., 1998 *(48)*	HC1	9	1	2	55	0.900	0.965	0.818	0.982	0.164	0.149
Ferris et al., 1998 *(29)*	HC1	5	5	53	106	0.500	0.667	0.086	0.955	0.343	0.059
Manos et al., 1999 *(49)*	HC2	58	7	326	582	0.892	0.641	0.151	0.988	0.395	0.067
Bergeron et al., 2000 *(50)*	HC2	10	2	38	61	0.833	0.616	0.208	0.968	0.432	0.108
Fait et al., 2000 *(51)*	HC1	48	8	5	165	0.857	0.971	0.906	0.954	0.235	0.248
Lin et al., 2000 *(52)*	HC2	27	0	12	35	1.000	0.745	0.692	1.000	0.527	0.365
Shlay et al., 2000 *(43)*	HC2	14	1	47	133	0.933	0.739	0.230	0.993	0.313	0.077
Morin et al., 2001 *(31)*	HC2	17	2	88	253	0.895	0.742	0.162	0.992	0.292	0.053
Rebello et al., 2001 *(53)*	HC2	18	3	13	41	0.857	0.759	0.581	0.932	0.413	0.280
Solomon et al., 2001 *(32)*	HC2	256	11	1050	984	0.959	0.484	0.196	0.989	0.568	0.116
Zielinski et al., 2001 *(54)*	HC2	11	1	63	138	0.917	0.687	0.149	0.993	0.347	0.056

*VT = ViraType; VP = ViraPap; HC = Hybrid Capture.

sensitivities (34.8% and 25.4%, respectively) and high specificities (96.8% and 99.2%, respectively).

Forest Plots

We created forest plots for the accuracy of all HPV tests combined (Fig. 1, upper panel) and for the Hybrid Capture II assay alone (Fig. 1, lower panel). Interstudy variation in sensitivity and specificity among all the HPV testing methods combined was large. The heterogeneity associated with the sensitivity of the Hybrid Capture II assay was much less than that for all HPV testing methods combined. Variation in the specificity of the Hybrid Capture II assay was reduced but was still substantial due to the low value observed in the ALTS study *(32)*.

The heterogeneity associated with sensitivity of repeat cytology at the cutoffs ASCUS or worse and LSIL or worse and with specificity at cutoff ASCUS was large, whereas the interstudy variation in specificity at the cutoff LSIL or worse was substantially smaller. Fig. 2 clearly shows that changing the threshold from ASCUS or worse to LSIL or worse increased the sensitivity values and decreased the specificity values.

Table 4. Triage of atypical squamous cells of undetermined significance (ASCUS) by repeat cytology at cutoffs of ASCUS or worse, low-grade squamous intraepithelial lesion (LSIL) or worse, and high-grade squamous intraepithelial lesion (HSIL) or worse for the detection of histologically confirmed cervical intraepithelial neoplasia grade II or worse (CIN2+): number of true- and false-positives and true- and false-negatives, accuracy parameters, and test positivity rate derived from different published studies*

Study	Type	True-positive	False-negative	False-positive	True-negative	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Test positivity rate
					Cutoff of ASCUS or worse					
Goff et al., 1993 *(46)*	CP	3	2	75	91	0.600	0.548	0.038	0.978	0.456
Slawson et al., 1994 *(47)*	CP	10	5	35	71	0.667	0.670	0.222	0.934	0.372
Cox et al., 1995† *(18)*	CP	11	4	NA	NA	0.733	NA	NA	NA	NA
Wright et al., 1995 *(19)*	CP	8	3	81	89	0.727	0.524	0.090	0.967	0.492
Ferris et al., 1998 *(29)*	CP	7	3	88	71	0.700	0.447	0.074	0.959	0.562
Manos et al., 1999 *(49)*	ThinPr	48	15	324	570	0.762	0.638	0.129	0.974	0.389
Bergeron et al., 2000 *(50)*	CP	8	4	28	71	0.667	0.717	0.222	0.947	0.324
Morin et al., 2001 *(31)*	CP	14	5	126	214	0.737	0.629	0.100	0.977	0.390
Solomon et al., 2001 *(32)*	ThinPr	227	40	1132	914	0.850	0.447	0.167	0.958	0.588
					Cutoff of LSIL or worse					
Goff et al., 1993 *(46)*	CP	2	3	6	160	0.400	0.964	0.250	0.982	0.047
Cox et al., 1995† *(18)*	CP	8	7	NA	NA	0.533	NA	NA	NA	NA
Wright et al., 1995 *(19)*	CP	6	5	22	148	0.545	0.871	0.214	0.967	0.155
Ferris et al., 1998 *(29)*	CP	2	8	23	136	0.200	0.855	0.080	0.944	0.148
Manos et al., 1999 *(49)*	ThinPr	25	38	52	842	0.397	0.942	0.325	0.957	0.080
Morin et al., 2001 *(31)*	CP	8	11	22	318	0.421	0.935	0.267	0.967	0.084
Solomon et al., 2001 *(32)*	ThinPr	158	109	452	1593	0.592	0.779	0.259	0.936	0.264
					Cutoff of HSIL or worse					
Manos et al., 1999 *(49)*	ThinPr	16	47	7	887	0.254	0.992	0.696	0.950	0.024
Solomon et al., 2001 *(32)*	ThinPr	93	174	67	2002	0.348	0.968	0.581	0.920	0.068

*CP = conventional Pap smear; ThinPr = smear prepared using ThinPrep liquid-based technology. NA = data not available.

†Cox *(18)* was not included for computation of specificity of repeat cytology because of unavailability of data.

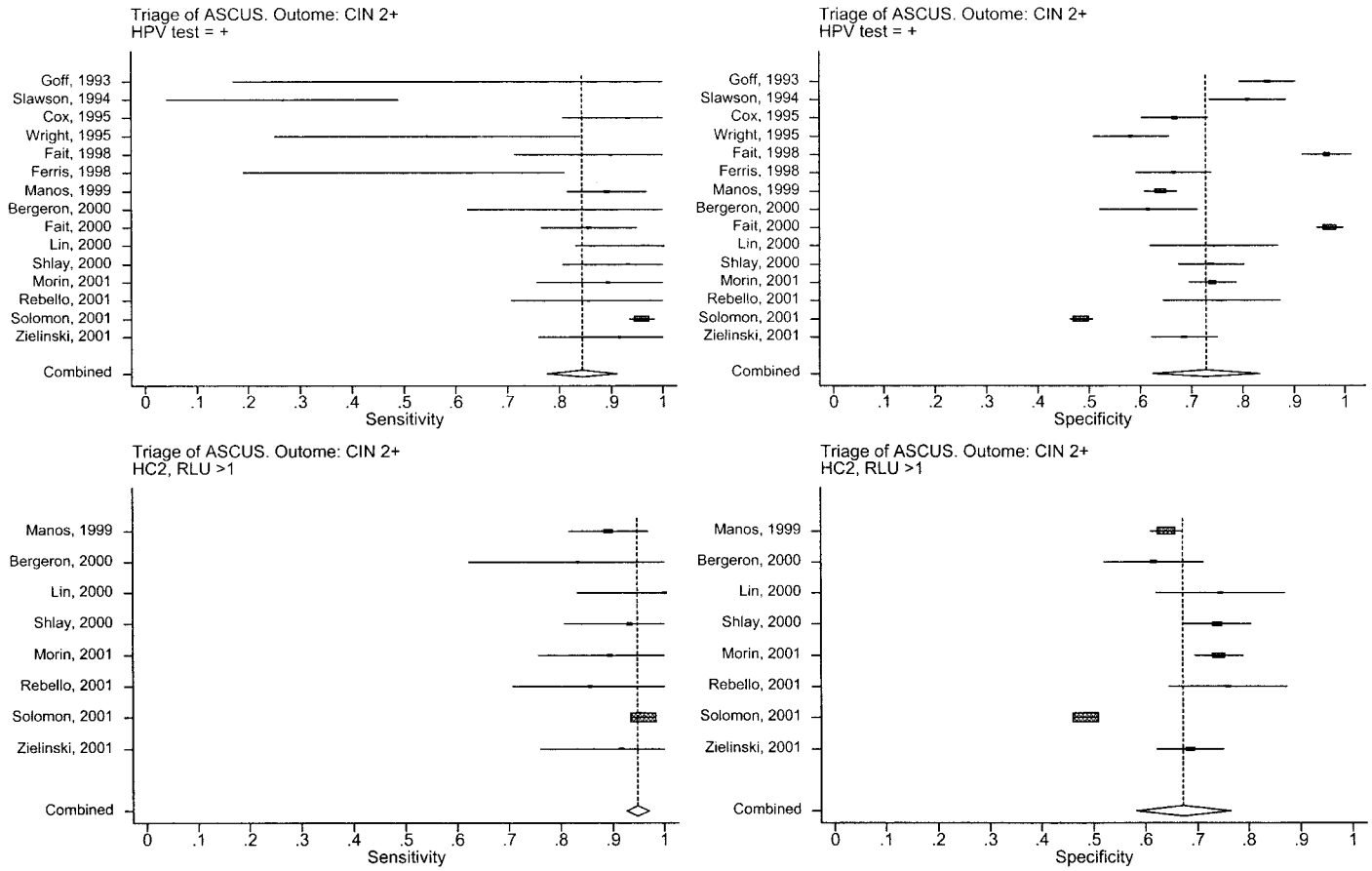


Fig. 1. Forest plots of the sensitivity and the specificity of human papillomavirus (HPV) DNA detection methods defined at the base level threshold for all 15 studies (**upper panel**) and for the eight studies in which the Hybrid Capture II assay was used (**lower panel**). The value of the plotted parameter is represented by a **rectangle**, the surface of which is proportional to the weight that the study contributes to the meta-analysis. The width of the horizontal line depicts the 95% confidence interval. The 95% confidence interval of the pooled estimate is displayed at the bottom as a **diamond**.

Meta-analysis

We computed pooled estimates of the sensitivity and specificity (diamonds in Figs. 1 and 2) and of predictive values and likelihood ratios (Table 5). Meta-analysis including all HPV tests yielded a pooled sensitivity of 84.4% (95% CI = 77.6% to 91.1%) and a specificity of 72.9% (95% CI = 62.5% to 83.3%). Consideration of just the eight studies *(31,32,43,49,50,52–54)* in which the Hybrid Capture II assay was used for the detection of high-risk HPV types yielded a sensitivity of 94.8% (95% CI = 92.7% to 96.9%) and a specificity of 67.3% (95% CI = 58.2% to 76.4%).

The sensitivity and specificity of repeat cytology at cutoff ASCUS or worse was 81.8% (95% CI = 73.5% to 84.3%) and 57.6% (95% CI = 49.5% to 65.7%), respectively. The sensitivity and specificity of repeat cytology at cutoff LSIL or worse were 45.7% (95% CI = 34.0% to 57.4%) and 89.1% (95% CI = 82.1% to 96.2%), respectively—substantially lower and higher than at a cutoff of ASCUS or worse.

The pooled ratio of the sensitivity of the Hybrid Capture II assay to the sensitivity of repeat cytology at cutoff ASCUS or worse, determined from four studies that evaluated both triage methods, was 1.16 (95% CI = 1.04 to 1.29), suggesting that the pooled sensitivity of the Hybrid Capture II assay was statistically significantly higher than that of repeat cytology at cutoff ASCUS or worse. The pooled specificity of the Hybrid Capture II assay was also higher, but the difference was not statistically significant (1.05, 95% CI = 0.96 to 1.15). Comparison of the Hybrid Capture II assay with repeat cytology at the respective cutoffs LSIL or worse and HSIL or worse yielded sensitivity ratios of 1.69 (95% CI = 1.54 to 1.85) and 2.80 (95% CI = 2.43 to 3.31), respectively, and specificity ratios of 0.71 (95% CI = 0.64 to 0.80) and 0.57 (95% CI = 0.44 to 0.74), respectively.

The negative likelihood ratio for the Hybrid Capture II assay was substantially lower than that for repeat cytology. The positive likelihood ratio for cytologic triage at cutoff LSIL or worse was substantially higher than that at cutoff ASCUS or worse because of its high specificity despite a low sensitivity. Both primary studies that provided data for cytologic triage at cutoff HSIL or worse *(32,49)* had extremely low sensitivity (meta-analysis data not shown).

The prevalence of CIN2+ ranged from 3% *(46)* to 36% *(52)*, with a pooled mean of 10.5% (95% CI = 7.9% to 13.1%). Studies in which women with repeated atypia were evaluated had the highest prevalence of CIN2+ *(52,53)*.

Influence of Study Characteristics

We determined the change in pooled sensitivity and specificity over technical and design variables for HPV DNA detection methods and repeat cytology at cutoff ASCUS or worse (Table 6). The sensitivity varied by viral test technique: it was low for

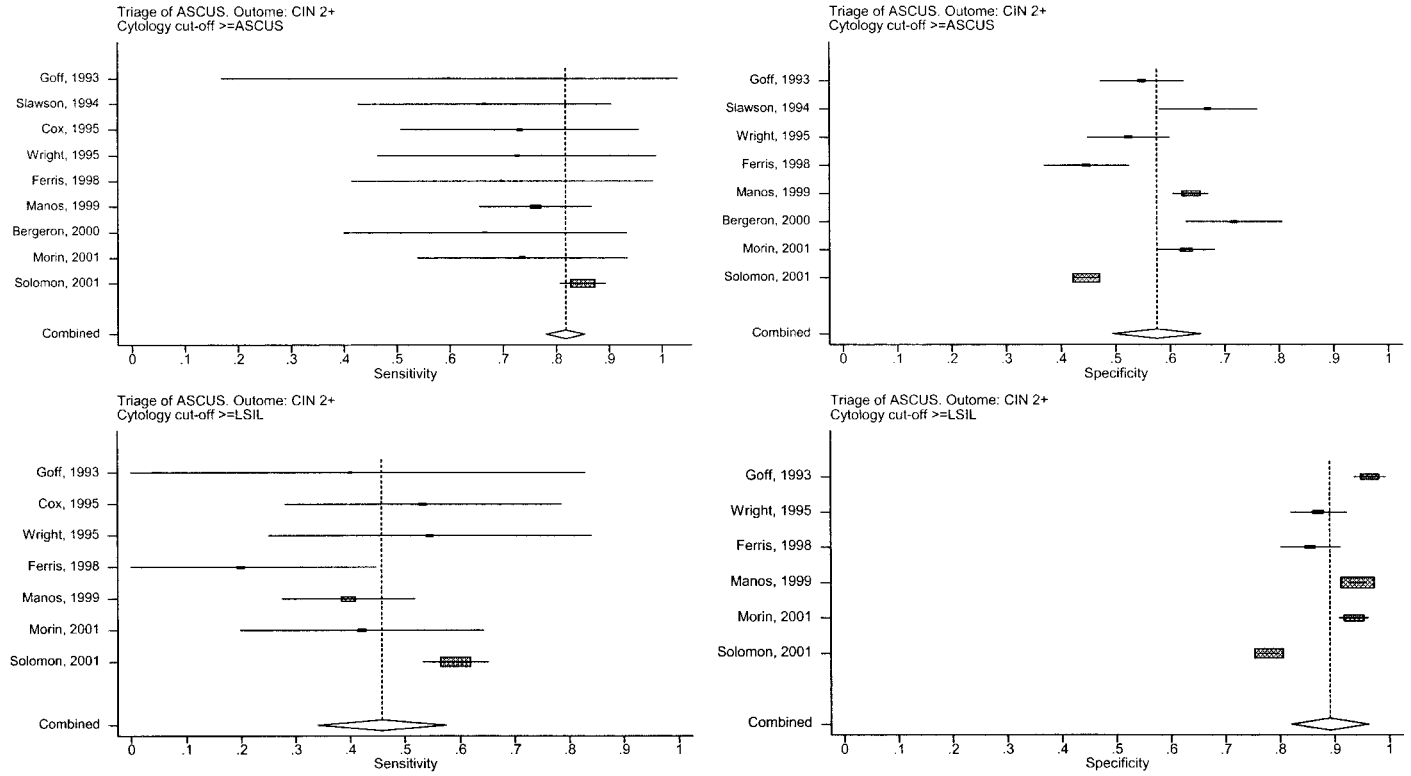


Fig. 2. Forest plots of the sensitivity and the specificity of repeat cytology defined at thresholds greater than or equal to atypical squamous cells of undetermined significance (≥ASCUS, **upper panel**) and greater than or equal to low-grade squamous intraepithelial lesion (≥LSIL, **lower panel**) for the presence of histologically confirmed cervical intraepithelial neoplasia grade II or worse (CIN2+).

Table 5. Summary of the meta-analyses of the accuracy of atypical squamous cells of undetermined significance (ASCUS) triage to predict the presence of cervical intraepithelial neoplasia grade II or worse (CIN2+): pooled estimates and confidence intervals, model of pooling, range (lowest to highest accuracy), and number of included studies*

Triage test	Test cutoff	Accuracy parameter	Model	Pooled estimate (95% CI)	Range	No. of studies
HPV DNA testing (all)	Positive signal	Sensitivity	R	0.844 (0.776 to 0.911)	0.267–1.000	15
		Specificity	R	0.729 (0.625 to 0.833)	0.484–0.971	15
		PPV	R	0.301 (0.203 to 0.398)	0.068–0.906	15
		NPV	R	0.985 (0.979 to 0.992)	0.887–1.000	15
		PLR	R	2.970 (2.380 to 3.690)	1.310–29.140	15
		NLR	R	0.240 (0.130 to 0.430)	0.010–0.660	14
HC2	RLU >1	Sensitivity	F	0.948 (0.927 to 0.969)	0.833–1.000	8
		Specificity	R	0.673 (0.582 to 0.764)	0.484–0.759	8
		PPV	R	0.264 (0.192 to 0.336)	0.149–0.906	8
		NPV	F	0.990 (0.985 to 0.994)	0.932–1.000	8
		PLR	R	2.820 (2.240 to 3.540)	1.860–3.920	8
		NLR	F	0.130 (0.090 to 0.190)	0.080–0.270	7
Repeat cytology	ASCUS	Sensitivity	F	0.818 (0.735 to 0.843)	0.600–0.850	9
		Specificity	R	0.576 (0.495 to 0.657)	0.447–0.717	8
		PPV	R	0.118 (0.076 to 0.159)	0.038–0.222	8
		NPV	F	0.967 (0.960 to 0.975)	0.934–0.987	8
		PLR	R	1.750 (1.490 to 2.060)	1.260–2.360	8
		NLR	F	0.390 (0.320 to 0.480)	0.340–0.730	8
	LSIL	Sensitivity	R	0.457 (0.340 to 0.574)	0.200–0.592	7
		Specificity	R	0.891 (0.821 to 0.962)	0.779–0.964	6
		PPV	R	0.232 (0.160 to 0.304)	0.080–0.325	6
		NPV	R	0.958 (0.943 to 0.974)	0.936–0.982	6
		PLR	R	4.440 (2.600 to 7.610)	1.380–11.070	6
		NLR	R	0.630 (0.520 to 0.770)	0.520–0.940	6

*CI = confidence interval; F = fixed model of pooling; HC = Hybrid Capture; HPV = human papillomavirus; LSIL = low-grade squamous intraepithelial lesion; NLR = negative likelihood ratio; NPV = negative predictive value; PLR = positive likelihood ratio; PPV = positive predictive value; R = random-effects model of pooling; RLU = relative light units.

Table 6. Variation of the sensitivity and specificity of human papillomavirus (HPV) DNA detection methods and repeat cytology by study characteristics*

Covariate	Sensitivity (95% CI)	Method†	Specificity (95% CI)	Method†
	HPV testing for an outcome of CIN2+			
HPV DNA detection method				
ViraPap, ViraType	0.381 (0.071 to 0.691)	R	0.836 (0.792 to 0.880)	F
HC1, HCT	0.788 (0.670 to 0.905)	R	0.789 (0.669 to 0.908)	R
HC2	0.948 (0.927 to 0.969)	F	0.673 (0.582 to 0.764)	R
Sampling device				
Dacron swab	0.635 (0.409 to 0.860)	R	0.756 (0.601 to 0.911)	R
Cone brush	0.919 (0.853 to 0.985)	F	0.701 (0.650 to 0.752)	R
Cotton applicator	0.900 (0.714 to 1.000)	B	0.965 (0.917 to 1.000)	B
Broom	0.936 (0.873 to 0.998)	R	0.562 (0.408 to 0.716)	R
Not documented	0.824 (0.542 to 1.000)	R	0.759 (0.609 to 0.928)	R
Transport medium				
ViraPap	0.570 (0.215 to 0.925)	R	0.682 (0.592 to 0.772)	R
STM	0.891 (0.839 to 0.944)	F	0.796 (0.689 to 0.902)	R
ThinPrep	0.936 (0.873 to 0.998)	R	0.562 (0.408 to 0.716)	R
Not documented	0.824 (0.542 to 1.000)	R	0.769 (0.609 to 0.928)	R
Year of publication				
≤1998	0.638 (0.395 to 0.881)	R	0.759 (0.638 to 0.880)	R
>1998	0.943 (0.923 to 0.964)	F	0.709 (0.564 to 0.854)	R
Quality review of histological outcome				
No	0.796 (0.696 to 0.896)	R	0.763 (0.680 to 0.847)	R
Yes	0.952 (0.930 to 0.975)	F	0.596 (0.470 to 0.721)	R
Blinding of outcome verification				
No	0.878 (0.811 to 0.944)	R	0.740 (0.578 to 0.902)	R
Yes	0.786 (0.624 to 0.949)	R	0.706 (0.650 to 0.762)	R
Clinical history: old CIN cases included				
No	0.920 (0.863 to 0.976)	R	0.658 (0.567 to 0.750)	R
Yes	0.754 (0.619 to 0.890)	R	0.790 (0.688 to 0.892)	R
Prevalence of CIN2+, %				
<10	0.859 (0.779 to 0.939)	R	0.698 (0.642 to 0.754)	R
≥10–<20	0.752 (0.487 to 1.000)	R	0.719 (0.442 to 0.996)	R
≥20	0.883 (0.812 to 0.953)	F	0.833 (0.658 to 1.000)	R
	Repeat cytology at cutoff ASCUS or worse for an outcome of CIN2+			
Sampling device				
Ayre and Cytobrush	0.712 (0.569 to 0.856)	F	0.619 (0.429 to 0.809)	R
Broom	0.820 (0.737 to 0.902)	R	0.542 (0.355 to 0.729)	R
Unknown	0.697 (0.569 to 0.825)	F	0.574 (0.482 to 0.665)	R
Year of publication				
≤1998	0.699 (0.579 to 0.819)	F	0.545 (0.460 to 0.630)	R
>1998	0.801 (0.728 to 0.873)	R	0.604 (0.476 to 0.733)	R
Preparation of smear				
Conventional	0.704 (0.608 to 0.799)	F	0.588 (0.513 to 0.663)	R
LBC	0.820 (0.737 to 0.902)	R	0.542 (0.355 to 0.729)	R
Quality review of histological outcome				
No	0.697 (0.591 to 0.803)	F	0.588 (0.513 to 0.660)	R
Yes	0.834 (0.795 to 0.873)	F	0.542 (0.355 to 0.729)	R
Blinding of outcome verification				
No	0.836 (0.795 to 0.876)	F	0.570 (0.467 to 0.673)	R
Yes	0.740 (0.656 to 0.825)	F	0.586 (0.463 to 0.708)	R
Clinical history, old CIN cases included				
No	0.828 (0.790 to 0.866)	F	0.541 (0.422 to 0.660)	R
Yes	0.677 (0.537 to 0.816)	F	0.612 (0.521 to 0.703)	R
Prevalence of CIN2+, %				
<10	0.742 (0.665 to 0.819)	F	0.563 (0.495 to 0.630)	R
≥10–<20	0.775 (0.636 to 0.914)	R	0.608 (0.411 to 0.804)	R

*CIN2+ = cervical intraepithelial neoplasia grade II or worse; ASCUS = atypical squamous cells of undetermined significance; HC1 = Hybrid Capture I assay; HCT = Hybrid Capture Tube test; HC2 = Hybrid Capture II assay; CI = confidence interval; STM = specimen transport medium; LBC = liquid-based cytology.

†A fixed (F) model was used for pooling. However, a random-effects pooling method (R) was chosen if the *P* value for the heterogeneity test >0.2, because the number of studies in the subgroups was often small. B = when only one study was concerned, confidence intervals were computed according to the binomial distribution.

ViraPap and ViraType (38.1%, 95% CI = 7.1% to 69.1%), higher for the Hybrid Capture I/Hybrid Capture Tube (78.8%, 95% CI = 67.0% to 90.5%), and even higher for the Hybrid Capture II assay (94.8%, 95% CI = 92.7% to 96.9%). The use of different HPV detection methods over time yielded a period effect for test sensitivity. The sensitivity of HPV DNA detection was also higher in studies in which a cervical broom or a conical brush was used for sample collection and in which the histologic

outcome was submitted to quality review. There was no difference in sensitivity or specificity among cytologic triage subgroups.

Differences in accuracy by age group could be assessed from only two studies *(33,43)* (data not shown). A statistically significant increase in specificity for triage with Hybrid Capture II and cytologic triage at a cutoff of ASCUS or worse and a statistically nonsignificant decrease in sensitivity for the Hybrid Capture II assay were observed with increasing age.

Publication Bias

We found a lower sensitivity for cytologic triage and a higher specificity for cytologic and virologic triage in smaller studies than in larger studies. This relation was, to a large extent, driven by the ALTS results *(32,33)*, which include younger women. However, no asymmetry was found when odds ratios of accuracy for HPV triage and repeat cytology were considered. We conclude that the selective publication of smaller studies did not affect our main conclusions.

Discussion

In this meta-analysis, we evaluated the accuracy of two triage tests used for women who had a previous equivocal Pap smear, focusing on the internal test validity parameters of sensitivity and specificity because of their more universal properties. We identified several sources of interstudy variability from the retrieved literature and identified potential biases that can affect the quality of the evaluation of test accuracy. Predictive values are dependent on local disease prevalence and therefore have limited generalizability. Likelihood ratios only have a relative value—for pooling data dependent on local prevalence—they are clinically less useful parameters than sensitivity and specificity parameters. For example, a high positive likelihood was found for triage at LSIL or worse due to a relatively high specificity (89%) despite a low sensitivity (46%).

Extent of Interstudy Heterogeneity

There was a wide range of sensitivities and specificities for cytologic triage. This was expected because the reproducibility of the cytologic interpretation of Pap smears is often reported as moderate to poor, especially for women with atypia or borderline disease *(68–70)*.

In the ALTS, only 32.4% of women with an original interpretation of ASCUS had a subsequent ASCUS result at enrollment, on average 2 months later *(71)*. In the other studies, the proportion of women with ASCUS at index and enrollment smears varied from 17% *(47)* to 41% *(29)*. The low reproducibility of cytology decreases the capacity to detect or to exclude intraepithelial disease and is likely to result in the variability in accuracy observed in the forest plots (Fig. 2). Accuracy varied even more for HPV triage methods than for cytology (Fig. 1), largely because of the use of different HPV DNA detection systems. It was therefore not recommended that they be combined into one pooled measure of HPV DNA testing. Nevertheless, it was useful to describe this variability among historical test systems, because we could clearly show that not all HPV DNA detection systems have the same accuracy. The sensitivity of the Hybrid Capture II assay showed low variability, reflecting the improved reliability of HPV DNA detection with the Hybrid Capture assay. Indeed, both the Hybrid Capture I and Hybrid Capture II assays have excellent reproducibility (unweighted kappa = 0.85 and 0.72, respectively) *(19,72)*.

Specificity for the Hybrid Capture II assay was consistent across studies, with the only exception being that it was considerably lower in the ALTS. A possible explanation is that the majority of women in the ALTS belong to younger age groups in whom HPV infection may be common in the absence of cervical lesions *(33,73)*.

Other HPV DNA Detection Methods

Certain authors applied additional HPV detection methods other than the Hybrid Capture assays. Morin et al. *(31)* used a polymerase chain reaction (PCR) amplifying a 450-base-pair sequence from the L1 region of the HPV genome, followed by a probe targeting 11 high-risk types. Compared with that of the Hybrid Capture II assay, the sensitivity of the PCR test to detect women with CIN2+ was equal (89.5% [95% CI = 66.9% to 98.7%]), but the specificity was substantially lower (59.0% [95% CI = 53.6% to 64.1%] for PCR versus 74.2% [95% CI = 69.2% to 78.8%] for the Hybrid Capture II assay). In a population of women with ASCUS or LSIL, Bergeron et al. *(50)* found a statistically nonsignificantly higher sensitivity for a PCR test that used MY9/11 HPV consensus primers (96%, 95% CI = 77% to 99%) than for the combined low- and high-risk probe of the Hybrid Capture II assay (86%, 95% CI = 65% to 97%), although the specificity for each test was similar (40% versus 41%). Bergeron et al. also used Southern blot hybridization to detect HPV, which was considerably more specific (65%, 95% CI = 60% to 71%) but also much less sensitive (46%, 95% CI = 24% to 68%) than PCR and the Hybrid Capture II assay. From Table 6, we conclude that certain HPV tests (ViraPap, ViraType, and Hybrid Capture I) cannot be recommended for triage of women with ASCUS. The performance of PCR-based systems is insufficiently documented in the ASCUS triage setting and needs further study.

Robustness of the Meta-analysis Results

This meta-analysis appears to corroborate the conclusions of the ALTS *(32,33,74)*. However, from the forest plots (Figs. 1 and 2), the results from the ALTS are rather extreme. Sensitivity values were 85.0% and 59.2% for triage at cutoffs ASCUS or worse and LSIL or worse, respectively, in the ALTS, but 81.8% (95% CI = 73.5% to 84.3%) and 45.7% (95% CI = 34.0% to 57.4%), respectively, in our meta-analysis. The specificity was 44.7% for repeat cytology at the threshold of ASCUS or worse in the ALTS and 57.6% (95% CI = 49.5% to 65.7%) in our meta-analysis. The sensitivity of the Hybrid Capture II assay was 95.9% in the ALTS and 94.8% (95% CI = 92.7% to 96.9%) in our pooled data, and its specificity was 48.4% in the ALTS and 67.3% (95% CI = 58.2% to 76.4%) in our pooled data.

To investigate whether the difference in sensitivity between repeat cytology and Hybrid Capture II assay was influenced by the ALTS results, we repeated the meta-analysis excluding the ALTS results. A meta-analysis considering only three studies *(31,49,50)* resulted in a pooled sensitivity ratio of 1.18 (95% CI = 1.03 to 1.36) and a pooled specificity ratio of 1.03 (95% CI = 0.89 to 1.19). Thus, inclusion of the ALTS data did not influence the pooled relative accuracy values.

Verification Bias

Because colposcopy was performed on every woman, with the exception of those in the ALTS, there was no verification bias in our analysis. In the ALTS, results from women in the HPV arm were not verified if the Hybrid Capture II assay was negative, if results from repeat cytology ranged from normal to LSIL, or if no suspect macroscopic lesions were observed. However, colposcopy was performed on all women in a second arm of the ALTS. Because the detection rates of histologically confirmed CIN2+ were 11.3% (95% CI = 9.5% to 13.2%) for women in the colposcopy arm and 11.7% (95% CI = 9.9% to 13.7%) for women in the HPV arm, we conclude that there is no evidence of verification bias in the ALTS that could have potentially influenced our results.

Validity of the Gold Standard

We used colposcopy and histology as the gold standard test. An imperfect gold standard can influence the estimation of the sensitivity and specificity of a triage test, which can be assessed by examining the variation of accuracy by disease prevalence *(75)*. However, the accuracy did not change substantially, even over wide ranges of disease prevalence (Table 6). The sensitivity was higher when the colposcopist was aware of the triage test results (for repeat cytology at cutoff ASCUS or worse: 81% if unmasked versus 73% if masked; for HPV triage: 88% if unmasked versus 79% if masked). Although this difference was not statistically significant, it suggests variability in the validity of the gold standard. The specificities were similar for masked and unmasked evaluations.

Low Specificity of All Triage Methods

The specificity of triage with the Hybrid Capture II assay or by repeat Pap smears at a low cytologic threshold was moderate to poor. Colposcopy of all triage-positive women generates considerable costs. Therefore, there is a need for more specific tests with high predictive value that allow for the identification of women at increased risk for cervical cancer. Nevertheless, a recent cost–effectiveness simulation study of alternative scenarios for ASCUS management *(76)* indicated that Hybrid Capture II is more effective and cost-effective than repeat cytology. HPV DNA testing can be done on residual fluid if liquid-based cytology is performed or if an additional sample is collected with the first smear, both of which reduce costs. In contrast, repeat cytology requires multiple visits *(77)*. When evaluating the cost–effectiveness of HPV-based triage, it should be noted that the specificity of the test is highly age-dependent and that cost may therefore vary depending on the age of the population to be screened.

Cross-sectional Outcome: Detection of CIN2+

Our analysis was restricted to studies documenting accuracy of triage methods for the presence of moderate dysplasia or worse (CIN2+). In a subsequent study, we will assess triage considering severe dysplasia and carcinoma *in situ* (CIN3) as the outcome. CIN3+ may be a more relevant endpoint because its potential to regress spontaneously is more limited than in CIN2 *(15,78,79)*.

Our meta-analysis was further restricted to cross-sectional outcomes in which sensitivity and specificity were measured against a simultaneously applied gold standard (colposcopically directed biopsy). Longitudinal studies are required to assess the possibility of detecting missed lesions by repeat cytology before invasive cancer occurs *(80)*. The ALTS group recently published their longitudinal trial outcomes, defined in terms of cumulative incidence of histologic CIN3+ over a period of 2 years, in which all subjects were followed at 6-month intervals and finally submitted to a colposcopy/histology check-up *(81)*. The longitudinal sensitivity of HPV DNA testing only at enrollment for cumulative CIN3+ was estimated to be 92.4% (95% CI = 88.7% to 95.2%). Testing ASCUS women only once with Hybrid Capture II required referral of 53.1% (95% CI = 51.4% to 54.8%) of women in the HPV triage arm. However, repeat cytology at every visit with a threshold of ASCUS would potentially detect 97.2% (95% CI = 94.1% to 100%) of the cumulative CIN3+ cases. This repeat cytology scenario would refer 73% (95% CI = 70.1% to 75.4%) of all women for colposcopy and would require multiple successive visits and cytologic examinations. Nevertheless, our meta-analysis of cross-sectional accuracy for CIN2+ is relevant because women with ASCUS require a follow-up decision and CIN2+ is the usual threshold for treatment.

Delineation of the Categories of Atypical Squamous Cells According to the 2001 Bethesda Reporting System

Our meta-analysis concerns women with a previous interpretation of ASCUS, such as that defined at the 1988 and 1991 Bethesda Workshops *(35,82)*. In 2001, a new version of the Bethesda Reporting System for Cervical Cytology was adopted that proposed a change in the subdivision of the global ASCUS class from three to two classes: ASC-US (atypical squamous cells of undetermined significance) and ASC-H (atypical squamous cells, HSIL cannot be ruled out) *(35,83)*. According to the new reporting system, the previous classification of atypical squamous cells favoring a benign reactive process (ASC-R) is now classified as negative for intraepithelial lesions or malignancy. The American Society for Colposcopy and Cervical Pathology has also recently updated its guidelines for the management of cervical cytologic abnormalities according to the new Bethesda terminology. Detection of high-risk HPV DNA was proposed as the first choice for triage of women with a liquid-based cytology result of ASC-US, whereas direct referral for colposcopic exploration is proposed as the optimal approach for women with ASC-H *(84,85)*. ASC-H is estimated to account for approximately 5%–10% of the ASC category *(83,86–88)*. In the ALTS, the prevalence of high-risk HPV types in the ASC-H group was 86%, and the proportion of women with ASC-H that contained CIN2+ was 41% *(87)*. HPV triage is therefore redundant for women with ASC-H.

It is not clear to what extent the elimination of ASC-R from the equivocal neoplastic changes will modify the conclusions of our meta-analysis. ASC-R account for approximately 20%–50% of all ASCUS results in U.S. women and a lower percentage in European women *(87–91)*. Because the prevalence of high-risk HPV types in ASC-R lesions is substantially lower than in other ASC lesions and because the positive predictive value of ASC-R for high-grade disease is minimal, HPV triage is probably not cost-effective for women with ASC-R results.

The accuracy of high-risk HPV DNA detection for triage of women with ASC results defined according to the 2001 Be-

thesda Reporting System guidelines is insufficiently documented. Nevertheless, we expect that restricting HPV triage to women with ASC-US will reduce the volume of HPV testing and considerably optimize its efficiency in the detection of women with CIN2+.

Recent Reviews and Meta-analyses

HPV triage has recently been reviewed *(28,92,93)*. Cuzick et al. *(28)* concluded that the available evidence did not support recommending widespread HPV testing in primary screening but that limited use, such as in management of women with borderline results, might be suggested. In the 2002 update of the French guidelines for management of women with cytologic abnormalities *(92)*, high-risk HPV DNA testing is recommended in triage of women with ASC-US. The Australian Medical Services Advisory Committee team reviewed literature regarding the management of women with low-grade epithelial abnormalities (LGEA) and concluded that there was insufficient evidence to support reimbursement for HPV triage of LGEA *(93)*. The authors remarked that international data did not match the particular Australian cytologic reporting system and that throughout the cytologic literature different thresholds for triage tests and outcomes were used, which made pooling data difficult *(93)*. However, no effort was made to pool data by separate test thresholds and histologic outcomes. The Australian LGEA category encompasses essentially low-grade squamous intraepithelial lesions. Our meta-analysis concerned the triage of women with ASCUS and not women with LSIL. However, in another systematic review *(34)* and also in the ALTS *(33,94,95)* it was found that HPV triage for LSIL had limited use because the test positivity rate was too high. Disparities between the Australian review and ours are likely the result of differences in the inclusion criteria for the index smear.

In conclusion, evidence is available indicating improved cross-sectional sensitivity of the Hybrid Capture II assay in comparison with the repeat Pap smear considered at cutoff ASCUS or worse for the outcome of high-grade CIN among women with equivocal cytologic results. The specificity of both triage methods is low. Cytologic triage of women with ASCUS that considers higher cytologic cutoffs yields unacceptably low sensitivity. We conclude that the Hybrid Capture II assay is a better triage method than repeat cytology for women with ASCUS.

References

(1) Franco EL, Rohan TE, editors. Cancer precursors. Epidemiology, detection and prevention. New York (NY): Springer-Verlag; 2002. p. 1–430.

(2) Kurman RJ, Henson DE, Herbst AL, Noller KL, Schiffman MH. Interim guidelines for management of abnormal cervical cytology. The 1992 National Cancer Institute Workshop. JAMA 1994;271:1866–9.

(3) Cox JT. HPV testing: is it useful in triage of minor Pap abnormalities? J Fam Pract 1998;46:121–4.

(4) Robertson JH, Woodend BE, Crozier EH, Hutchinson J. Risk of cervical cancer associated with mild dyskaryosis. BMJ 1988;297:18–21.

(5) Coleman D, Day N, Douglas G, Farmery E, Lynge E, Philip J, et al. European Guidelines for Quality Assurance in Cervical Cancer Screening. Europe against cancer programme. Eur J Cancer 1993;29A Suppl 4:S1–38.

(6) Flannelly G, Anderson D, Kitchener HC, Mann EM, Campbell M, Fisher P, et al. Management of women with mild and moderate cervical dyskaryosis. BMJ 1994;308:1399–403.

(7) Arbyn M, Albertyn G, De Groof V. Consensus on follow-up recommendations in cervical cancer screening (in Dutch). Tijdschr voor Geneeskunde 1996;52:709–15.

(8) Noumoff JS. Atypia in cervical cytology as a risk factor for intraepithelial neoplasia. Am J Obstet Gynecol 1987;156:628–31.

(9) Richart RM. Causes and management of cervical intraepithelial neoplasia. Cancer 1987;60(8 Suppl):1951–9.

(10) Richart RM, Wright TC Jr. Controversies in the management of low-grade cervical intraepithelial neoplasia. Cancer 1993;71(4 Suppl):1413–21.

(11) Soutter WP. Management of women with mild dyskaryosis. Immediate referral to colposcopy is safer. BMJ 1994;309:591–2.

(12) Ferenczy A. Viral testing for genital human papillomavirus infections: recent progress and clinical potentials. Int J Gynecol Cancer 1995;5:321–8.

(13) Wilkinson C, Jones JM, McBride J. Anxiety caused by abnormal result of cervical smear test: a controlled trial. BMJ 1990;300:440.

(14) Narod SA, Thompson DW, Jain M, Wall C, Green LM, Miller AB. Dysplasia and the natural history of cervical cancer: early results of the Toronto Cohort Study. Eur J Cancer 1991;27:1411–6.

(15) Ostor AG. Natural history of cervical intraepithelial neoplasia: a critical review. Int J Gynecol Pathol 1993;12:186–92.

(16) Melnikow J, Nuovo J, Willan AR, Chan BK, Howell LP. Natural history of cervical squamous intraepithelial lesions: a meta-analysis. Obstet Gynecol 1998;92:727–35.

(17) Murdoch JB, Grimshaw RN, Morgan PR, Monaghan JM. The impact of loop diathermy on management of early invasive cervical cancer. Int J Gynecol Cancer 1992;2:129–33.

(18) Cox JT, Lorincz AT, Schiffman MH, Sherman ME, Cullen A, Kurman RJ. Human papillomavirus testing by hybrid capture appears to be useful in triaging women with a cytologic diagnosis of atypical squamous cells of undetermined significance. Am J Obstet Gynecol 1995;172:946–54.

(19) Wright TC, Sun XW, Koulos J. Comparison of management algorithms for the evaluation of women with low-grade cytologic abnormalities. Obstet Gynecol 1995;85:202–10.

(20) Kinney WK, Manos MM, Hurley LB, Ransley JE. Where's the high-grade cervical neoplasia? The importance of minimally abnormal Papanicolaou diagnoses. Obstet Gynecol 1998;91:973–6.

(21) zur Hausen H. Molecular pathogenesis of cancer of the cervix and its causation by specific human papillomavirus types. Curr Top Microbiol Immunol 1994;186:131–56.

(22) Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International Biological Study on Cervical Cancer (IBSCC) Study Group. J Natl Cancer Inst 1995;87:796–802.

(23) Munoz N. Human papilloma virus and cervical cancer: epidemiological evidence. In: Franco E, Monsonego J, editors. New developments in cervical cancer screening and prevention. Oxford (UK): Blackwell Science; 1997. p. 3–13.

(24) Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999;189:12–9.

(25) Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV. The causal relation between human papillomavirus and cervical cancer. J Clin Pathol 2002; 55:244–65.

(26) Kaufman RH, Adam E. Does typing of human papillomavirus assist in the triage of women with repeated low-grade, cervical cytologic abnormalities? Gynecol Oncol 1998;70:317–8.

(27) Kaufman RH, Adam E. Is human papillomavirus testing of value in clinical practice? Am J Obstet Gynecol 1999;180:1049–53.

(28) Cuzick J, Sasieni P, Davies P, Adams J, Normand C, Frater A, et al. A systematic review of the role of human papillomavirus testing within a cervical screening programme. Health Technol Assess 1999;3:i–iv, 1–196.

(29) Ferris DG, Wright TC Jr, Litaker MS, Richart RM, Lorincz AT, Sun XW, et al. Triage of women with ASCUS and LSIL on Pap smear reports: management by repeat Pap smear, HPV DNA testing, or colposcopy? J Fam Pract 1998;46:125–34.

(30) Ferris DG, Wright TC Jr, Litaker MS, Richart RM, Lorincz AT, Sun XW, et al. Comparison of two tests for detecting carcinogenic HPV in women with Papanicolaou smear reports of ASCUS and LSIL. J Fam Pract 1998; 46:136–41.

(31) Morin C, Bairati I, Bouchard C, Fortier M, Roy M, Moore L, et al. Managing atypical squamous cells of undetermined significance in Papanicolaou smears. J Reprod Med 2001;46:799–805.
(32) Solomon D, Schiffman M, Tarone R; ALTS Study group. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. J Natl Cancer Inst 2001;93:293–9.
(33) Sherman ME, Schiffman M, Cox JT; Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study Group. Effects of age and human papilloma viral load on colposcopy triage: data from the randomized Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study (ALTS). J Natl Cancer Inst 2002;94:102–7.
(34) Arbyn M, Buntinx F, Van Ranst M, Cortiñas Abrahantes J. Triage of women with atypical or low-grade cytological abnormalities of the cervix by HPV testing: systematic review and meta-analysis. Brussels (Belgium): Institute of Public Health/Epidemiology Reports No. 2001–019[D/2001/2505/34]. 2002. p. 1–244.
(35) Luff RD. The Bethesda System for reporting cervical/vaginal cytologic diagnoses: report of the 1991 Bethesda workshop. The Bethesda System Editorial Committee. Hum Pathol 1992;23:719–21.
(36) Richart RM. Cervical intraepithelial neoplasia. Pathol Annu 1973;8:301–28.
(37) Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. Stat Med 1988;7:889–94.
(38) Light RJ, Singer JD, Willet JB. The visual presentation and interpretation of meta-analysis. In: Cooper H, editor. The handbook of research synthesis. New York (NY): Russell Sage Foundation; 1994. p. 439–54.
(39) Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Methods for meta-analysis in medical research. Chichester (UK): John Wiley & Sons, Ltd.; 2000. p. 1–317.
(40) Cochran WG. The combination of estimates from different experiments. Biometrics 1954;10:101–29.
(41) Dersimonian R, Laird NM. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
(42) Sharp S, Sterne J. Meta-analysis. Stata Technical Bulletin 1997;38:9–14.
(43) Shlay JC, Dunn T, Byers T, Baron AE, Douglas JM Jr. Prediction of cervical intraepithelial neoplasia grade 2–3 using risk assessment and human papillomavirus testing in women with atypia on Papanicolaou smears. Obstet Gynecol 2000;96:410–6.
(44) Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.
(45) Steichen TJ. Tests for publication bias in meta-analysis. Stata Technical Bulletin 1998;7:125–33.
(46) Goff BA, Muntz HG, Bell DA, Wertheim I, Rice LW. Human papillomavirus typing in patients with Papanicolaou smears showing squamous atypia. Gynecol Oncol 1993;48:384–8.
(47) Slawson DC, Bennett JH, Simon LJ, Herman JM. Should all women with cervical atypia be referred for colposcopy: a HARNET study. Harrisburgh Area Research Network. J Fam Pract 1994;38:387–92.
(48) Fait G, Daniel Y, Kupferminc MJ, Lessing JB, Niv J, Bar-Am A. Does typing of human papillomavirus assist in the triage of women with repeated low-grade, cervical cytologic abnormalities? Gynecol Oncol 1998;70:319–22.
(49) Manos MM, Kinney WK, Hurley LB, Sherman ME, Shieh-Ngai J, Kurman RJ, et al. Identifying women with cervical neoplasia: using human papillomavirus DNA testing for equivocal Papanicolaou results. JAMA 1999;281:1605–10.
(50) Bergeron C, Jeannel D, Poveda J, Cassonnet P, Orth G. Human papillomavirus testing in women with mild cytologic atypia. Obstet Gynecol 2000;95:821–7.
(51) Fait G, Kupferminc MJ, Daniel Y, Geva E, Ron IG, Lessing JB, et al. Contribution of human papillomavirus testing by hybrid capture in the triage of women with repeated abnormal pap smears before colposcopy referral. Gynecol Oncol 2000;79:177–80.
(52) Lin CT, Tseng CJ, Lai CH, Hsueh S, Huang HJ, Law KS. High-risk HPV DNA detection by Hybrid Capture II. An adjunctive test for mildly abnormal cytologic smears in women $>$ or $=$ 50 years of age. J Reprod Med 2000;45:345–50.
(53) Rebello G, Hallam N, Smart G, Farquharson D, McCafferty J. Human papillomavirus testing and the management of women with mildly abnormal cervical smears: an observational study. BMJ 2001;322:893–4.
(54) Zielinski DG, Snijders PJ, Rozendaal L, Voorhorst FJ, Runsink AP, de Schipper FA, et al. High-risk HPV testing in women with borderline and mild dyskaryosis: long-term follow-up data and clinical relevance. J Pathol 2001;195:300–6.
(55) Cox JT, Schiffman MH, Winzelberg AJ, Patterson JM. An evaluation of human papillomavirus testing as part of referral to colposcopy clinics. Obstet Gynecol 1992;80:389–95.
(56) Hatch KD, Schneider A, Abdel-Nour MW. An evaluation of human papillomavirus testing for intermediate- and high-risk types as triage before colposcopy. Am J Obstet Gynecol 1995;172:1150–5; discussion 1155–7.
(57) Herrington CS, Evans MF, Hallam NF, Charnock FM, Gray W, McGee JD. Human papillomavirus status in the prediction of high-grade cervical intraepithelial neoplasia in patients with persistent low-grade cervical cytological abnormalities. Br J Cancer 1995;71:206–9.
(58) Ferenczy A, Franco E, Arseneau J, Wright TC, Richart RM. Diagnostic performance of Hybrid Capture human papillomavirus deoxyribonucleic acid assay combined with liquid-based cytologic study. Am J Obstet Gynecol 1996;175:651–6.
(59) Hall S, Lorincz A, Shah F, Sherman ME, Abbas F, Paull G, et al. Human papillomavirus DNA detection in cervical specimens by hybrid capture: correlation with cytologic and histologic diagnoses of squamous intraepithelial lesions of the cervix. Gynecol Oncol 1996;62:353–9.
(60) Bollen LJ, Tjong-A-Hung SP, Van der Velden J, Brouwer K, Mol BW, ten Kate FJ, et al. Human papillomavirus deoxyribonucleic acid detection in mildly or moderately dysplastic smears: a possible method for selecting patients for colposcopy. Am J Obstet Gynecol 1997;177:548–53.
(61) Kaufman RH, Adam E, Icenogle J, Reeves WC. Human papillomavirus testing as triage for atypical squamous cells of undetermined significance and low-grade squamous intraepithelial lesions: sensitivity, specificity, and cost-effectiveness. Am J Obstet Gynecol 1997;177:930–6.
(62) Kaufman RH, Adam E, Icenogle J, Lawson H, Lee N, Reeves KO, et al. Relevance of human papillomavirus screening in management of cervical intraepithelial neoplasia. Am J Obstet Gynecol 1997;176:87–92.
(63) Adam E, Kaufman RH, Berkova Z, Icenogle J, Reeves WC. Is human papillomavirus testing an effective triage method for detection of high-grade (grade 2 or 3) cervical intraepithelial neoplasia? Am J Obstet Gynecol 1998;178:1235–44.
(64) Nyirjesy I, Billingsley FS, Forman MR. Evaluation of atypical and low-grade cervical cytology in private practice. Obstet Gynecol 1998;92:601–7.
(65) Lytwyn A, Sellors JW, Mahony JB, Daya D, Chapman W, Ellis N, et al. Comparison of human papillomavirus DNA testing and repeat Papanicolaou test in women with low-grade cervical cytologic abnormalities: a randomized trial. HPV Effectiveness in Lowgrade Paps (HELP) Study No. 1 Group. CMAJ 2000;163:701–7.
(66) Mould TA, Singer A, Gallivan S. Quantitative detection of oncogenic HPV DNA using hybrid capture to triage borderline and mildly dyskaryotic Papanicolaou smears. Eur J Gynaecol Oncol 2000;21:245–8.
(67) Lorincz AT. Methods of DNA hybridisation and their clinical applicability to human papillomavirus detection. In: Franco E, Monsonego J, editors. New developments in cervical cancer screening and prevention. Oxford (UK): Blackwell Science; 1997. p. 325–37.
(68) O'Sullivan JP, Ismail SM, Barnes WS, Deery AR, Gradwell E, Harvey JA, et al. Interobserver variation in the diagnosis and grading of dyskaryosis in cervical smears: specialist cytopathologists compared with non-specialists. J Clin Pathol 1994;47:515–8.
(69) Sherman ME, Schiffman MH, Lorincz AT, Manos MM, Scott DR, Kuman RJ, et al. Toward objective quality assurance in cervical cytopathology. Correlation of cytopathologic diagnoses with detection of high-risk human papillomavirus types. Am J Clin Pathol 1994;102:182–7.
(70) Gatscha RM, Abadi M, Babore S, Chhieng D, Miller MJ, Saigo PE. Smears diagnosed as ASCUS: interobserver variation and follow-up. Diagn Cytopathol 2001;25:138–40.
(71) Stoler MH, Schiffman M; Atypical Squamous Cells of Undetermined Significance-Low-grade Squamous Intraepithelial Lesion Triage Study (ALTS) Group. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA 2001;285:1500–5.

(72) Castle PE, Lorincz AT, Mielzynska-Lohnas I, Scott DR, Glass AG, Sherman ME, et al. Results of human papillomavirus DNA testing with the hybrid capture 2 assay are reproducible. J Clin Microbiol 2002;40:1088–90.
(73) Woodman CB, Collins S, Winter H, Bailey A, Ellis J, Prior P, et al. Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. Lancet 2001;357:1831–6.
(74) Schiffman M, Adrianza ME. ASCUS-LSIL Triage Study. Design, methods and characteristics of trial participants. Acta Cytol 2000;44:726–42.
(75) Boyko EJ. Re: "Meta-analysis of Pap test accuracy". Am J Epidemiol 1996;143:406–7.
(76) Kim JJ, Wright TC, Goldie SJ. Cost-effectiveness of alternative triage strategies for atypical squamous cells of undetermined significance. JAMA 2002;287:2382–90.
(77) Wright TC Jr, Lorincz A, Ferris DG, Richart RM, Ferenczy A, Mielzynska I, et al. Reflex human papillomavirus deoxyribonucleic acid testing in women with abnormal Papanicolaou smears. Am J Obstet Gynecol 1998; 178:962–6.
(78) Morrison SA. Screening in chronic disease. In: Kelsey JL, Marmot MG, Stolley PD, Vessey MP, editors. Series Monogr Epidemiol Biostat. Vol 19. 2nd ed. Oxford (UK): Oxford University Press; 1992. p. 1–254.
(79) Schiffman M, Kjaer SK. Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia. J Natl Cancer Inst Monogr 2003; 31:14–9.
(80) Sherman ME, Wang SS, Tarone R, Rich L, Schiffman M. Histopathologic extent of cervical intraepithelial neoplasia 3 lesions in the atypical squamous cells of undetermined significance low-grade squamous intraepithelial lesion triage study: implications for subject safety and lead-time bias. Cancer Epidemiol Biomarkers Prev 2003;12:372–9.
(81) ASCUS-LSIL Triage Study (ALTS) Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. Am J Obstet Gynecol 2003;188:1383–92.
(82) The 1988 Bethesda System for reporting cervical/vaginal cytologic diagnoses. National Cancer Institute Workshop. JAMA 1989;262:931–4.
(83) Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA 2002;287:2114–9.
(84) Wright TC Jr, Cox JT, Massad LS, Twiggs LB, Wilkinson EJ; ASCCP-Sponsored Consensus Conference. 2001 Consensus Guidelines for the management of women with cervical cytological abnormalities. JAMA 2002;287:2120–9.
(85) Stoler MH. New Bethesda terminology and evidence-based management guidelines for cervical cytology findings. JAMA 2002;287:2140–1.
(86) Sherman ME, Tabbara SO, Scott DR, Kurman RJ, Glass AG, Manos MM, et al. "ASCUS, rule out HSIL": cytologic features, histologic correlates, and human papillomavirus detection. Mod Pathol 1999;12:335–42.
(87) Sherman ME, Solomon D, Schiffman M; ASCUS LSIL Triage Study Group. Qualification of ASCUS. A comparison of equivocal LSIL and equivocal HSIL cervical cytology in the ASCUS LSIL Triage Study. Am J Clin Pathol 2001;116:386–94.
(88) Quddus MR, Sung CJ, Steinhoff MM, Lauchlan SC, Singer DB, Hutchinson ML. Atypical squamous metaplastic cells: reproducibility, outcome, and diagnostic features on ThinPrep Pap test. Cancer 2001;93:16–22.
(89) Quddus MR, Zhang S, Sung CJ, Liu F, Neves T, Struminsky J, et al. Utility of HPV DNA detection in thin-layer, liquid-based tests with atypical squamous metaplasia. Acta Cytol 2002;46:808–12.
(90) Hughes SA, Sun D, Gibson C, Bellerose B, Rushing L, Chen H, et al. Managing atypical squamous cells of undetermined significance (ASCUS): human papillomavirus testing, ASCUS subtyping, or follow-up cytology? Am J Obstet Gynecol 2002;186:396–403.
(91) Scott DR, Hagmar B, Maddox P, Hjerpe A, Dillner J, Cuzick J, et al. Use of human papillomavirus DNA testing to compare equivocal cervical cytologic interpretations in the United States, Scandinavia, and the United Kingdom. Cancer 2002;96:14–20.
(92) Agence Nationale d'Accréditation et d'Evaluation en Santé (ANAES). Management of a patient with an abnormal cervical smear. Service Recommandations et Références Professionnelles et Service Evaluation Economique. Paris (France): ANAES; 2002. p. 1–92.
(93) Medical Services Advisory Committee (MSAC). Human papillomavirus testing in women with cytological prediction of low-grade abnormality. MSAC reference 12b, Canberra (Australia): 2003. p. 1–97.
(94) Human papillomavirus testing for triage of women with cytologic evidence of low-grade squamous intraepithelial lesions: baseline data from a randomized trial. The Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesions Triage Study (ALTS) Group. J Natl Cancer Inst 2000;92:397–402.
(95) ASCUS-LSIL Triage Study (ALTS) Group. A randomized trial on the management of low-grade squamous intraepithelial lesion cytology interpretations. Am J Obstet Gynecol 2003;188:1393–400.

NOTES

Supported by the Europe Against Cancer program of the European Commission (Directorate of SANCO, Luxembourg), and the DWTC/SSTC (Federal Services for Scientific, Cultural and Technical Affairs of the Federal Government, Belgium). We are grateful to Bernadette Claus, Marisa Abarca, and Jessica Timmermans (of the Scientific Institute of Public Health) for their assistance in retrieval of literature references, and to Francis Capet and Ralph Crott for critical comments.

Manuscript received August 11, 2003; revised December 17, 2003; accepted December 23, 2003.

Vaccine 30S (2012) F88–F99

Contents lists available at SciVerse ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Review

Evidence Regarding Human Papillomavirus Testing in Secondary Prevention of Cervical Cancer

Marc Arbyn[a,b,*], Guglielmo Ronco[c], Ahti Anttila[d], Chris J.L.M. Meijer[e], Mario Poljak[f], Gina Ogilvie[g], George Koliopoulos[h], Pontus Naucler[i], Rengaswamy Sankaranarayanan[j], Julian Peto[k]

[a] *Unit of Cancer Epidemiology, Scientific Institute of Public Health, Brussels, Belgium*
[b] *European Cancer Network, IARC, Lyon, France*
[c] *Unit of Cancer Epidemiology, Centro per la prevenzione Oncologica, Turin, Italy*
[d] *Mass Screening Registry, Finnish Cancer Registry, Helsinki, Finland*
[e] *Department of Pathology, VU University Medical Centre, Amsterdam, The Netherlands*
[f] *Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia*
[g] *Division of STI/HIV Prevention and Control, BC Centre for Disease Control, Vancouver, Canada*
[h] *Department of Obstetrics and Gynecology, University Hospital of Ioannina, Ioannina, Greece*
[i] *Department of Infectious Diseases; Department of Microbiology, Tumor & Cell Biology, Karolinska University Hospital, Stockholm, Sweden*
[j] *Screening Group, International Agency for Research on Cancer, Lyon, France*
[k] *London School of Hygiene and Tropical Medicine, London, UK*

ARTICLE INFO

Article history:
Received 31 May 2012
Received in revised form 20 June 2012
Accepted 22 June 2012

Keywords:
HPV
Cervical cancer
Screening
Triage
ASC-US
LSIL
Treatment
Surveillance after treatment
Meta-analysis
Systematic review
Diagnostic test accuracy

ABSTRACT

More than ever, clinicians need regularly updated reviews given the continuously increasing amount of new information regarding innovative cervical cancer prevention methods. A summary is given from recent meta-analyses and systematic reviews on 3 possible clinical applications of human papillomavirus (HPV) testing: triage of women with equivocal or low-grade cytologic abnormalities; prediction of the therapeutic outcome after treatment of cervical intraepithelial neoplasia (CIN) lesions, and last not but not least, primary screening for cervical cancer and pre-cancer. Consistent evidence is available indicating that HPV-triage with the Hybrid Capture® 2 assay (Qiagen Gaithersburg, Inc., MD, USA [previously Digene Corp.] (HC2) is more accurate (higher sensitivity, similar specificity) than repeat cytology to triage women with equivocal Pap smear results. Several other tests show at least similar accuracy but mRNA testing with the APTIMA® (Gen-Probe Inc., San Diego, CA, USA) test is similarly sensitive but more specific compared to HC2. In triage of low-grade squamous intraepithelial lesions (LSIL), HC2 is more sensitive but its specificity is substantially lower compared to repeat cytology. The APTIMA® test is more specific than HC2 without showing a loss in sensitivity. Identification of DNA of HPV types 16 and/or 18, or RNA from the five most carcinogenic HPV types allow selecting women at highest risk for CIN3+ but the sensitivity and negative predictive value of these markers are lower than full-range high-risk HPV (hrHPV) testing. After conservative treatment of cervical pre-cancer, HPV testing picks up more quickly, with higher sensitivity and not lower specificity, residual or recurrent high-grade CIN than follow-up cytology. Primary screening for hrHPV generally detects more CIN2, CIN3 or cancer compared to cytology at cut-off atypical squamous cells of undetermined significance (ASC-US) or LSIL, but is less specific. Combined HPV and cytology screening provides a further small gain in sensitivity at the expense of a considerable loss in specificity if positive by either test is referred to colposcopy, in comparison with HPV testing only. Randomised trials and follow-up of cohort studies consistently demonstrate a significantly lower cumulative incidence of CIN3+ and even of cancer, in women aged 30 years or older, who were at enrollment hrHPV DNA negative compared to those who were cytologically negative. The difference in cumulative risk of CIN3+ or cancer for double negative (cytology & HPV) versus only HPV-negative women is small. HC2, GP5+/6+ PCR (polymerase chain reaction), cobas® 4800 PCR (Roche Molecular Systems Inc., Alameda, CA, USA) and Real Time PCR (Abbott Molecular, Des Plaines, IL, USA) can be considered as clinically validated for use in primary screening. The loss in specificity associated with primary HPV-based screening can be compensated by appropriate algorithms involving reflex cytology and/or HPV genotyping for HPV16 or 18. There exists a substantial evidence base to support that HPV testing is advantageous both in triage of women with equivocal abnormal cytology, in surveillance after treatment of CIN lesions and in primary screening of women aged 30 years or older. However, the possible advantages offered by HPV-based screening require a well organised program with good compliance with screening and triage policies.

* Corresponding author. Tel.: +0032 2 6425021; fax: +0032 2 6425410.
E-mail address: marc.arbyn@wiv-isp.be (M. Arbyn).

0264-410X/$ – see front matter © 2012 Marc Arbyn.. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.vaccine.2012.06.095

This article forms part of a special supplement entitled "*Comprehensive Control of HPV Infections and Related Diseases*" Vaccine Volume 30, Supplement 5, 2012.

© 2012 Marc Arbyn.. Published by Elsevier Ltd. All rights reserved.

1. Introduction

The recognition of the strong causal relationship between persistent cervical infection with high-risk human papillomavirus (HPV) types and occurrence of cervical cancer [1] has led to the development of a series of HPV DNA or RNA tests. Detection of high-risk (hr) HPV DNA is considered to be potentially useful in three clinical applications: (1) as a triage test to select women, whose cytology is equivocal or mildly abnormal, needing referral for diagnosis and treatment; (2) as a follow-up test for women treated for high-grade cervical intraepithelial neoplasia (CIN) with local ablative or excisional therapy to predict cure or failure of treatment; and (3) as a primary screening test, solely or in combination with cervical cytology to detect cervical precancer and to rule it out in the predominantly healthy population.

In this chapter, we will update and extend previously conducted meta-analyses and systematic reviews which synthesize current knowledge on the performance of HPV testing in each of these three clinical applications. In particular, attention is given on the evidence regarding HPV-based primary screening as a new paradigm of cervical cancer prevention and on the identification of HPV assays, which fulfil minimal requirements, allowing use in primary cervical cancer screening.

2. Material and methods

For the purpose of this paper, previous meta-analyses on the absolute and relative performance of HPV- and cytology-based testing of cervical specimens were extended and updated [2]. New HPV-related assays, in addition to the Hybrid Capture® 2 (Qiagen Gaithersburg, Inc., MD, USA [previously Digene Corp.] (HC2), were evaluated as well, including both other HPV DNA assays and E6/E7 mRNA essays. Three major clinical applications were distinguished: (1) triage of women with a cervical cytology of atypical squamous cells of undetermined significance (ASC-US) or low-grade squamous intraepithelial lesions (LSIL); (2) follow-up of women after treatment of a high-grade lesion with the purpose to predict cure or failure; and (3) primary cervical cancer screening. The search strings, consulted bibliographic sources and the last date of literature retrieval are presented in Appendix 1 (web table).

2.1. Triage of women with minor abnormal cytology

The current review evaluated the accuracy of the HC2 assay for underlying CIN2+ and CIN3+ among women with ASC-US or LSIL cervical cytology. In studies where a repeat Pap smear was taken as well, the relative sensitivity and specificity of HC2 versus repeat cytology were evaluated, considering ASC-US or worse as a positive repeat cytology. Reports were included only when verification with a reference standard was available based on colposcopy followed by biopsies for all tested women. The outcome was based on the histology result, when available, accepting negative colposcopy, despite its known less-than-perfect sensitivity, as sufficient ascertainment for absence of the target disease. We also evaluated the accuracy of other HPV tests, as well as their relative accuracy compared to HC2.

2.2. Follow-up after treatment of high-grade CIN

Studies were selected if (1) women were treated for histologically-confirmed CIN2 or CIN3; (2) women had a cytology and HPV DNA test between 3 and 9 months after treatment; (3) there was follow-up during at least 18 months with colposcopy and biopsy for all women or in case of a positive HPV or cytology test. Treatment failure was defined as follow-up histologic diagnosis of CIN2+.

2.3. Primary screening

Criteria for inclusion of reports have been published previously [2,3]. Two types of study design were considered: (1) cross-sectional studies where women were submitted to concomitant testing with cervical cytology (conventional or liquid), a HPV DNA assay and, optionally, other screening tests and (2) randomised clinical trials where women were assigned to cytology, HPV testing or combined testing. In the assessment of absolute sensitivity and specificity, we distinguished three situations: (1) all cases were verified with a reference standard, (2) only screen test positive cases were verified and the assumption was made that none of the women being negative for all tests had underlying CIN2+ and (3) studies in which a random sample of women being negative for all tests was also submitted to verification. For the evaluation of relative sensitivity, we considered the ratio of absolute sensitivities including intra-arm comparisons of randomised trials with combined cytology and HPV testing and the ratio of the detection rates of CIN2+ from the inter-arm comparison of randomised controlled trials. Studies were selected only when the participating women were representative of the general population.

Particular attention was given to the pooling of published aggregated data regarding the baseline and longitudinal outcomes from randomised trials, as well as the cumulative risk for CIN3 and cancer according to the baseline HPV and cytology status observed from non-randomised cohorts.

The review also addressed the question of which high-risk (hrHPV) DNA assays fulfill clinical equivalency criteria recently established to allow claims for use in cervical cancer screening. The candidate test should demonstrate a relative sensitivity and specificity compared to a validated hrHPV DNA test (HC2 or GP5+/6+ polymerase chain reaction [PCR]) of $\geq$0.90 and $\geq$0.98, respectively [4]. A representative set of samples (minimally 60 CIN2+ cases, 800 $\leq$CIN1 cases) derived from a population-based screening cohort should be selected, as specified by Meijer CJLM *et al.* [4]. Moreover, a high reproducibility (lower 95% confidence bound $\geq$87%) should be reached.

3. Results

3.1. Triage of women with minor abnormal cytology

3.1.1. ASC-US

In the 39 retrieved studies (web table liststudies.xls-ASCUStriage), the pooled sensitivity of HC2 was 90.4% (95% confidence interval [CI]: 88.1–92.3%) and 93.7% (95% CI: 90.4–95.9%), whereas the pooled specificity was 58.3% (95% CI: 53.6–62.9%) and 52.3% (95% CI: 45.7–58.7%) for predicting presence or absence of CIN2+ or CIN3+, respectively (Table 1). In 10 of the 39 studies, a repeat smear was also taken. In these 10

Table 1
Summary of meta-analyses on the test performance of HPV DNA testing in three possible clinical applications: triage of minor cytologic abnormalities (ASC-US or LSIL), prediction of residual or recurrent CIN2+ after treatment and primary cervical cancer screening. Sensitivity and specificity (pooled estimate, range [minimum and maximum observed value]) to detect histologically confirmed CIN2+ or CIN3+, pooled test positivity rate, and prevalence of CIN.

Application	Test	Test cut-off	Outcome	Studies	Sensitivity		Specificity		Test positivity rate	Prevalence
					pooled estimate (95% CI)	range (%)	pooled estimate (95% CI)	range (%)	pooled estimate (95% CI)	pooled estimate (95% CI)
Triage ASC-US	HC2	1 pg/mL	CIN2+	39	90.4 (88.1–92.3)	60–100	58.3 (53.6–62.9)	27–79	48.2 (43.7–52.6)	12.0 (10.2–13.7)
			CIN3+	17	93.7 (90.4–95.9)	75–100	52.3 (45.7–58.7)	33–70		6.8 (5.0–8.6)
Triage LSIL	HC2	1 pg/mL	CIN2+	24	95.4 (94.0–96.5)	80–100	27.8 (23.8–32.1)	16–58	76.9 (73.5–80.2)	21.3 (17.7–24.9)
			CIN3+	14	96.4 (90.5–98.7)	76–100	23.7(19.4–28.7)	15–47		8.9 (6.6–11.2)
Prediction treatment failure	HC2/PCR	Diverse	Recurrent CIN2+	15	92.9 (85.0–96.8)	54–100	81.3 (74.7–86.5)	57–100	31.9 (28.8–42.1)	8.4 (6.5–10.4)
Primary screening	HC2	1 pg/mL	CIN2+	31	90.4 (88.0–92.8)	50–100	88.5 (87.0–90.0)	61–95	13.2 (11.4–14.9)	2.2 (1.8–2.6)
			CIN2+	14[a]	96.3 (94.5–98.1)	69–100	91.4 (89.8–92.9)	86–96	9.9 (8.1–11.8)	1.5 (1.1–1.9)
			CIN3+	22	95.3 (93.3–97.3)	62–100	89.0 (87.2–90.8)	82–95		1.1 (0.8–1.3)
	HC2 & cytology	1 pg/mL or ASC-US+	CIN2+	13	94.2 (90.8–97.6)	63–100	87.7 (85.0–90.3)	69–94	13.8 (10.9–16.8)	1.3 (1.0–1.6)
			CIN2+	8[a]	99.8 (99.0–100)	69–100	88.8 (85.5–92.1)	80–93	12.5 (8.7–16.2)	0.7 (0.5–0.9)
	GP5+/6+	+ signal	CIN2+	3	94.5 (94.2–96.9)	94–100	94.8 (93.3–96.3)	86–96	8.0 (5.6–10.4)	2.4 (1.0–3.7)
	Cervista®	+ signal	CIN3+	1	95.0 (91.5–98.6)	-	90.3 (89.7–90.9)	-	11.1 (10.4–11.8)	1.6 (1.4–1.9)
	APTIMA®	+ signal	CIN3+	2	97.4 (91.4–100)	92–100	90.4 (89.7–91.1)	91–92	10.2 (9.5–10.9)	0.7 (0.6–0.7)
	cobas® 4800	+ signal	CIN3+	1	92.0 (88.8–95.2)	-	92.0 (91.7–92.3)	-	8.6 (8.3–8.8)	0.6 (0.5–0.8)
	careHPV™	1 pg/mL	CIN3+	1	87.0 (73.2–100)	-	86.1 (84.7–87.5)	-	14.6 (13.2–16.1)	1.0 (0.6–1.4)

[a] Restricted to studies conducted in North-America or Europe.
ASC-US: Atypical squamous cells of undetermined significance; CI: Confidence interval; CIN: Cervical intraepithelial neoplasia; HC2: Hybrid Capture® 2; LSIL: Low-grade squamous epithelial lesion; PCR: Polymerase chain reaction.

Fig. 1. Meta-analysis of the sensitivity and specificity of virologic or cytologic surveillance after treatment of high-grade CIN to predict treatment failure (defined as residual or recurrent CIN2 or worse within 24 months). Orange: hrHPV testing with HC2 or PCR; Grey: cytology. Small hollow symbols represent individual studies. The large filled symbols correspond with the summary accuracy estimates for testing around 6 months after treatment with HPV (orange circle) or cytology (grey square), respectively. The full curve represents the HSROC curve and the interrupted line delineates the 95% confidence ellipse around the summary accurate estimates. CIN: Cervical intraepithelial neoplasia; HC2: Hybrid Capture® 2; hrHPV: High-risk human papillomavirus; HSROC curve: Hierarchical summary receiver-operating characteristic curve; PCR: Polymerase chain reaction.

studies, the pooled sensitivity of HC2 was significantly higher than that of repeat cytology at cut-off ASC-US+ to detect CIN2+ (relative sensitivity: 1.27 (95% CI: 1.16–1.39)). The pooled specificity of both triage methods did not differ significantly from each other (relative specificity: 0.99, 95% CI: 0.97–1.03). HC2 also detected more CIN3+ than repeat cytology (relative sensitivity: 1.14, 95% CI: 1.06–1.22). Other HPV tests were evaluated for use in triage of ASC-US in recent studies and the pooled absolute accuracy, as well as the relative accuracy compared to the HC2 assay for CIN2+, are shown in Table 2. In general, subtle differences in sensitivity and specificity are observed between most hrHPV DNA tests or high-risk panels of full-range HPV genotyping tests and the HC2. However, from a recent meta-analysis, it was concluded that testing for viral RNA of 14 hrHPV types (APTIMA® [Gen-Probe Inc., San Diego, CA, USA]) was as sensitive but more specific for identifying CIN2+ compared to HC2 (specificity ratio of 1.19 (95% CI: 1.08–1.37)) [5]. Moreover, the PreTect™ HPV-Proofer (developed by NorChip/BioMérieux Inc., Klokkarstua, Hurum, Norway) (detecting RNA of HPV16, 18, 31, 33 and 45) and tests identifying DNA of HPV16 and/or 18 show significantly lower sensitivity but higher specificity compared to HC2.

3.1.2. LSIL

In the 24 retrieved studies (web table liststudies.xls-LSILtriage), the pooled absolute sensitivity of HC2 was 95.4% (95% CI: 94.0–96.5%) and 96.4% (95% CI: 90.5–98.7%), whereas the pooled specificity was 27.8% (95% CI: 23.8–32.1%) and 23.7% (95% CI: 19.4–28.7%) for the outcomes of CIN2+ and CIN3+, respectively (see Table 1). In six of 24 studies allowing assessment of the accuracy of HC2 in triage of LSIL, a repeat smear was also taken. In these six studies, the pooled sensitivity of HC2 was significantly higher than that of repeat cytology at cut-off of ASC-US+ (ratio of 1.23; 95% CI: 1.06–1.4). However the specificity of HC2 testing was substantially and significantly lower (ratio of 0.66; 95% CI: 0.58–0.75).

Table 2
Left: absolute pooled sensitivity and specificity of HPV DNA or RNA testing in triage of women with ASC-US or LSIL to find underlying CIN2+. Right: relative sensitivity and specificity compared to HC2.

		Absolute accuracy			Relative accuracy compared to HC2		
Triage group	Test	Number of studies	Pooled absolute sensitivity (95% CI)	Pooled absolute specificity (95% CI)	Number of studies	Pooled relative sensitivity (95% CI)	Pooled relative specificity (95% CI)
ASC-US	HC2	39	90.4 (88.1–92.3)	58.3 (53.6–62.9)	-	-	-
	Amplicor	5	90.5 (86.4–94.6)	45.2 (32.5–58.0)	4	0.98 (0.92–1.05)	0.87 (0.80–0.95)
	Abbott RT PCR	4	94.7 (90.2–99.2)	39.5 (26.4–52.6)	4	0.97 (0.90–1.04)	1.18 (0.99–1.41)
	Linear Array	9	93.8 (90.2–97.5)	46.2 (37.9–54.5)	5	1.02 (0.99–1.06)	0.90 (0.79–1.03)
	Papillocheck	1	96.1 (92.4–99.9)	60.2 (51.0–69.4)	1	1.01 (0.95–1.07)	1.44 (1.10–1.89)
	Cervista	3	95.9 (93.1–98.7)	49.6 (34.6–64.6)	1	0.98 (0.95–1.02)	1.15 (1.06–1.24)
	Cobas-4800	2	89.9 (84.3–95.6)	69.2 (65.7–72.7)	1	1.03 (0.92–1.16)	0.99 (0.95–1.04)
	Pretect HPV-Proofer	5	80.7 (75.4–86.0)	79.2 (66.6–91.8)	3	0.79 (0.71–0.88)	1.76 (1.55–1.99)
	APTIMA	7	95.7 (91.5–97.2)	56.4 (44.7–67.5)	7	1.01 (0.97–1.06)	1.19 (1.08–1.31)
	HPV16	16	54.0 (50.2–57.9)	87.0 (83.9–90.1)	7	0.60 (0.55–0.65)	1.59 (1.35–1.87)
	HPV16/18	17	58.1 (58.2–63.0)	83.4 (80.1–86.8)	8	0.65 (0.60–0.72)	1.67 (1.35–2.06)
LSIL	HC2	24	95.4 (94.0–96.5)	27.8 (23.8–32.1)	-	-	-
	Amplicor	3	93.9 (84.3–100)	28.4 (16.8–39.9)	2	0.98 (0.90–1.07)	0.80 (0.65–0.98)
	Abbott RT PCR	3	96.5 (93.6–99.4)	31.9 (18.6–45.3)	3	0.97 (0.94–1.00)	1.23 (0.96–1.58)
	Linear Array	6	99.0 (97.7–100)	28.1 (22.5–33.6)	5	1.00 (0.96–1.04)	1.03 (0.84–1.26)
	Cervista	1	96.8 (94.0–99.6)	46.7 (40.6–52.8)	-	-	-
	Papillocheck	2	94.6 (79.4–100)	32.1 (19.9–44.4)	2	1.05 (0.93–1.19)	0.90 (0.54–1.51)
	Pretect HPV-Proofer	5	75.6 (70.8–80.3)	76.6 (70.0–83.1)	5	0.75 (0.67–0.84)	2.79 (2.10–3.71)
	APTIMA	7	91.0 (85.2–94.7)	42.5 (33.3–52.3)	7	0.96 (0.92–1.03)	1.37 (1.22–1.54)
	HPV16	14	51.1 (46.5–55.7)	80.5 (76.7–84.3)	10	0.55 (0.50–0.61)	2.48 (2.04–3.01)
	HPV16/18	14	57.8 (54.0–61.7)	75.0 (70.8–79.3)	10	0.59 (0.54–0.64)	2.34 (1.96–2.79)

ASC-US: Atypical squamous cells of undetermined significance; CIN: Cervical intraepithelial lesion; HC2: Hybrid Capture® 2; LSIL: Low-grade squamous intraepithelial lesion.

As for HC2, the sensitivity of most other general hrHPV DNA tests is high (>90% for high-grade CIN) in LSIL triage, whereas the specificity is low (up to 30% lower compared to ASC-US triage) (see lower panel of Table 2). However, testing for viral RNA with APTIMA® is significantly more specific than HC2 (ratio of 1.37; 95% CI: 1.22–1.54) without losing sensitivity (ratio of 0.96 for CIN2+ and 0.98 for CIN3+ with 95% CIs including unity for both outcomes) [5]. Again, as with ASC-US triage, genotyping for a limited number of types is much more specific, but at the expense of a substantial loss in sensitivity.

3.2. Follow-up after treatment of high-grade CIN

Fifteen studies could be included (eight with HC2 and seven with PCR) that contained data on the accuracy of hrHPV DNA testing and cytology to predict residual or recurrent CIN2+ after treatment (web table liststudies.xls-postTrFU). The pooled sensitivity was 93% (95% CI: 85–97%) for HPV and 72% (95% CI: 66–78%) for cytology (see Figs. 1 and 2 and Table 1). The pooled specificities were 81% (95% CI: 74–86%) and 84% (95% CI: 80–87%), for HPV and cytology, respectively. HPV testing was significantly more sensitive (ratio of 1.25; 95%

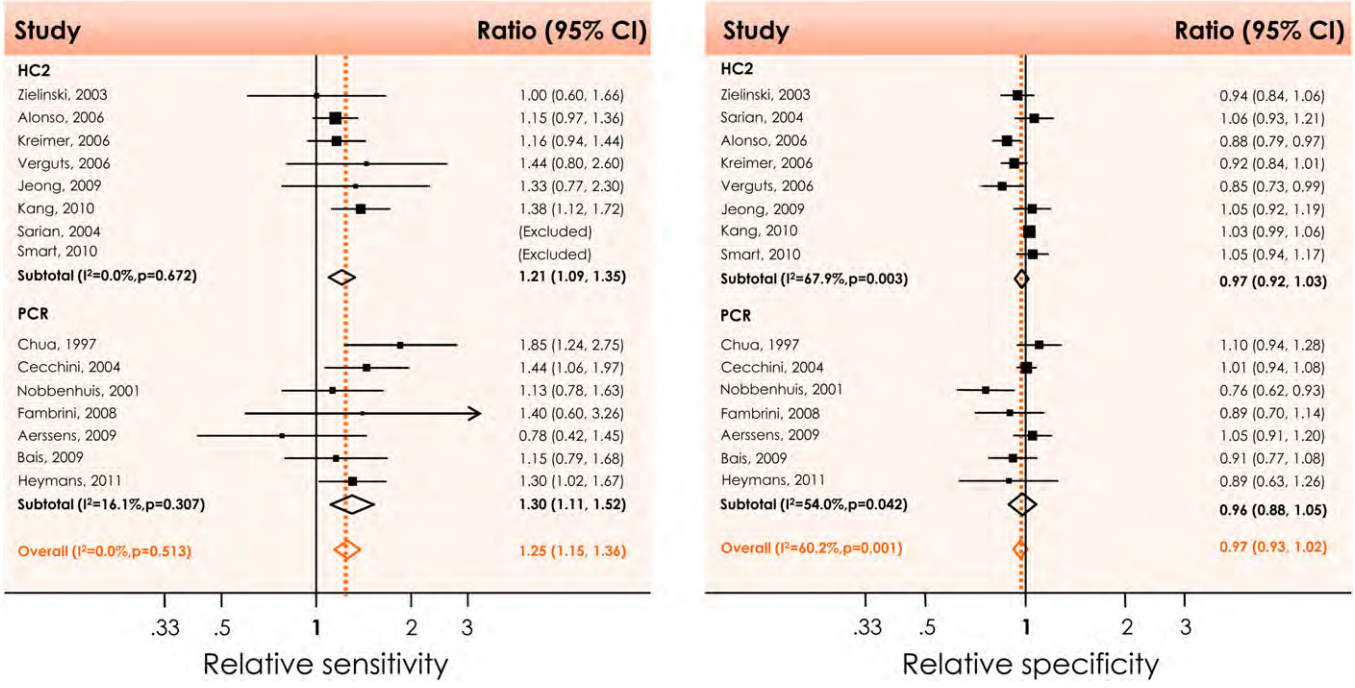


Fig. 2. Relative sensitivity (left) and specificity (right) of high-risk HPV DNA testing with HC2 or PCR compared to cytology after treatment of high-grade to predict therapeutic failure (residual of recurrent CIN2 or worse). See web table liststudies.xls-postTrFU for publication information. CI: Confidence interval; HC2: Hybrid Capture® 2; I^2: Percentage of total variation across studies due to heterogeneity; p: Test for inter-study heterogeneity; PCR: Polymerase chain reaction.

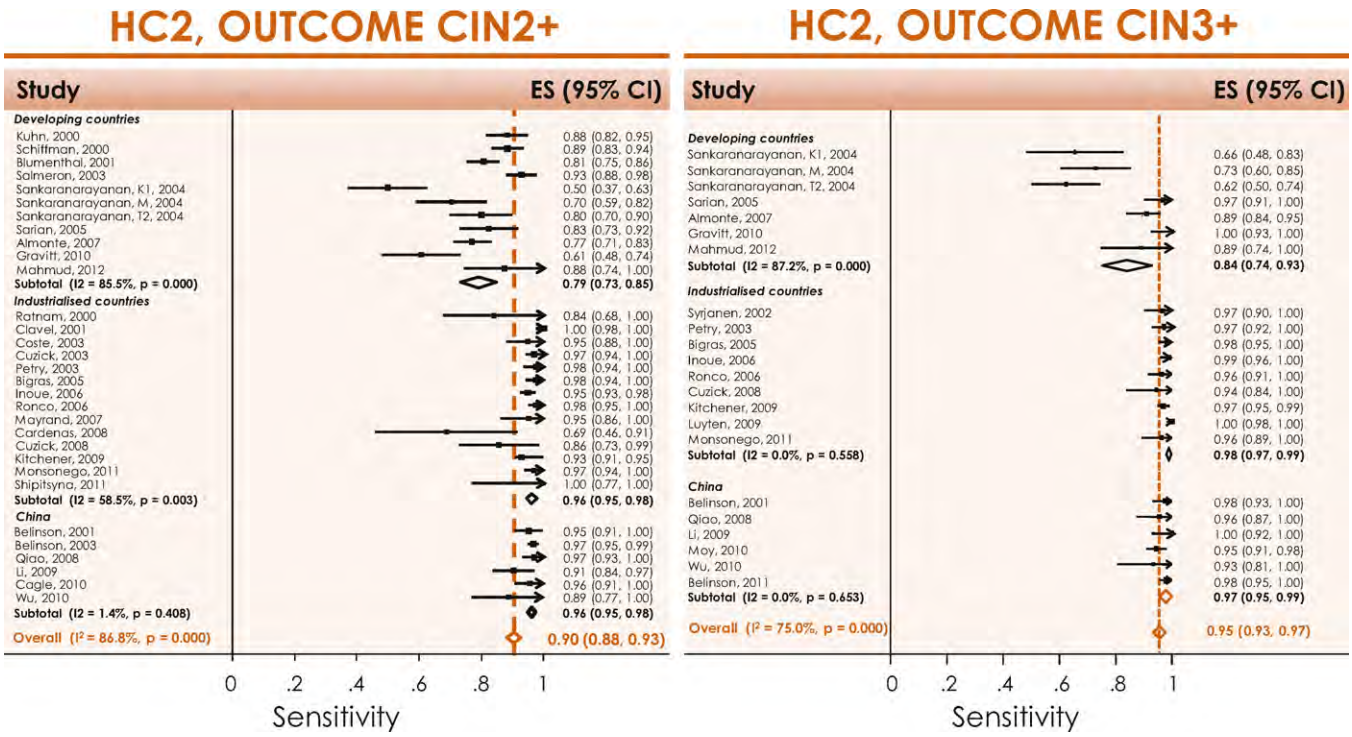


Fig. 3. Meta-analysis of the sensitivity of HC2 as a primary screening test to detect CIN2+ (left) or CIN3+ (right) in developing countries, industrialised countries and China. It shows clearly that the sensitivity is very heterogeneous in developing countries (probably due to quality of gold standard), much less heterogeneous in industrialised countries and not heterogeneous in China (improved gold standard verification). See web table liststudies.xls-screening for publication information. CIN: Cervical intraepithelial neoplasia; ES: Estimate of sensitivity; HC2: Hybrid Capture® 2; I^2: Percentage of total variation across studies due to heterogeneity; p: Test for inter-study heterogeneity.

CI: 1.15–1.36) but not less specific (ratio of 0.97; 95% CI: 0.93–1.02) compared to cytology (see Fig. 2). There was no heterogeneity by HPV testing system.

Combined testing with cytology and HPV was not significantly more sensitive but significantly less specific than HPV alone (relative sensitivity 1.07, 95% CI: 0.97–1.16; relative specificity 0.91, 95% CI: 0.88–0.94).

3.3. *Primary screening for cervical (pre-)cancer*

3.3.1. *Cross-sectional accuracy*

The updated meta-analysis of the cross-sectional accuracy of HPV tests contains data from 49 studies, among which 8 are randomised trials (web table liststudies.xls-screening). In the large majority of studies, the HC2 or GP5+/6+ PCR were used. Since the last couple of years data were published on the performance of HPV-based screening with other HPV DNA or RNA detection systems (Cervista® [Hologic, Madison, WI USA], APTIMA®, cobas® 4800 [Roche Molecular Systems Inc., Alameda, CA, USA], careHPV™ [Qiagen, Hilden, Germany], Abbott RT PCR [Abbott Molecular, Des Plaines, IL, USA], PapilloCheck® [Greiner Bio-One, Frickenhausen, Germany] and others) (Poljak M *et al.*, Vaccine, this issue [6]).

3.3.1.1. Absolute cross-sectional accuracy of HPV testing. Overall, the sensitivity of HC2 (at standard conditions: relative light unit [RLU]≥1) for detecting underlying CIN2+ and CIN3+ was 90% (95% CI: 88–93%) and 95% (95% CI: 93–97%), respectively (Fig. 3 and Table 1). The heterogeneity in the sensitivity was very large in studies conducted in developing countries (probably due to the variability in the quality of colposcopy and histology verification) [2]. The inter-study variation was substantially reduced in studies conducted in industrialised countries and was even non-significant in Chinese studies, in which an improved gold standard with multiple (random) biopsies was applied (Fig. 3).

In European and North American studies, the pooled sensitivity for CIN2+ was 96% (95% CI: 95–98%), whereas the pooled specificity was 91% (95% CI: 90–93%). The accuracy values of HC2 for CIN3+ were similar to those for CIN2+. Ten percent (95% CI: 8–12%) of the screened population was hrHPV-positive and 1.5% (95% CI: 1.1–1.9%) and 1.1% (95% CI: 0.8–1.3%) had CIN2+ or CIN3+, respectively.

The sensitivity of other HPV assays evaluated in primary screening was consistently high for detection of CIN3+ (in general > 92%, except careHPV™, at 87%). The specificity of these other tests varied between 86% and 92%.

3.3.1.2. Relative cross-sectional accuracy. The sensitivity of HC2 (at standard cut-off = 1 pg/mL) was on average 23–43% higher than cytology, depending on the disease threshold (CIN2+ or CIN3+) or the cytologic cut-off (ASC-US or LSIL) (see Table 3). The specificity of HC2 for excluding CIN2+ was significantly lower than cytology: ratio of 0.97 (95% CI: 0.96–0.98) and 0.92 (95% CI: 0.90–0.94), considering the cut-offs ASC-US+ or LSIL+, respectively.

The combination of cytology with HC2 was, on average, 42% (95% CI: 36–48%) and 33% (95% CI: 29–37%) higher for the detection of CIN2+ or CIN 3+, respectively, than cytology alone (at cut-off ASC-US+), whereas the specificity was 6% (95% CI: 6–7%) and 8% (95% CI: 7–9%) lower. In contrast, adding cytology to the HC2 increased the average sensitivity with only 5% (95% CI: 4–7%) for CIN2+ and 2% (95% CI: 1–3%) for CIN3+ compared to HPV testing alone; however, this resulted in a significant loss of specificity (ratios of 0.95; 95% CI: 0.94–0.96 and 0.93; 95% CI: 0.92–0.95, respectively).

The number of studies evaluating other tests than HC2 is still rather small (Fig. 4) [6]. Most of them showed a significantly higher sensitivity than cytology to detect high-grade CIN (ratios varying between 1.22 and 1.77), with the exception of certain older PCR systems targeting L1 sequences (Table 3). The specificity of most HPV assays was lower than cytology, although the difference was not always significant. PreTect™ and certain older L1 PCR systems showed a specificity similar to cytology.

Table 3
Relative accuracy of virologic versus cytologic screening or of combined screening versus testing with one test in order to find underlying CIN2 or CIN3 or worse.

Comparison	Outcome	Relative sensitivity (95% CI)		Range	Relative specificity (95% CI)		Range	No. of studies[a]
HC2/cyto (ASC-US+)	CIN2+	1.23	(1.15–1.31)	0.91–2.93	0.97	(0.96–0.98)	0.86–1.10	28/25[b]
HC2/cyto (ASC-US+)[c]		1.37	(1.22–1.54)	1.06–2.25	0.97	(0.96–0.98)	0.93–1.00	12/10[b]
HC2/cyto (LSIL+)		1.40	(1.27–1.54)	1.09–2.37	0.92	(0.90–0.94)	0.67–1.03	20/19[b]
HC2/cyto (ASC-US/LSIL+)		1.27	(1.18–1.36)	0.91–2.93	0.96	(0.94–0.97)	0.67–1.10	33/30[b]
HC2/cyto (ASC-US+)	CIN3+	1.27	(1.12–1.44)	0.97–2.63	0.97	(0.96–0.99)	0.88–1.10	20/18[b]
HC2/cyto (ASC-US+)[c]		1.43	(1.15–1.77)	1.01–2.12	0.97	(0.96–0.98)	0.93–1.00	8/6[b]
HC2/cyto (LSIL+)		1.36	(1.21–1.53)	0.97–2.32	0.93	(0.91–0.96)	0.84–1.03	13/12[b]
Cyto (ASC-US+) & HC2/Cyto (ASC-US+)	CIN2+	1.42	(1.36–1.48)	1.06–2.30	0.94	(0.93–0.94)	0.89–0.96	13
Cyto (ASC-US+) & HC2/Cyto (ASC-US+)	CIN3+	1.33	(1.29–1.37)	1.02–2.18	0.92	(0.91–0.93)	0.85–0.96	10/9
Cyto (ASC-US+) & HC2/HC2[d]	CIN2+	1.05	(1.04–1.07)	1.00–1.19	0.95	(0.94–0.96)	0.81–0.99	10
Cyto (ASC-US+) & HC2/HC2[d]	CIN3+	1.02	(1.01–1.03)	1.04–1.04	0.93	(0.92–0.95)	0.81–0.99	6
GP5+/6+/cyto (ASC-US+)	CIN2+	1.33	(1.13–1.55)	0.75–1.50	0.94	(0.85–1.03)	0.89–0.99	3/2
Cobas-4800/cyto (ASC-US+)	CIN2+	1.77	(1.55–1.89)	-	0.97	(0.97–0.97)	-	1
Other L1-based PCR/cyto (ASC-US+)	CIN2+	1.04	(0.87–1.25)	0.75–1.50	0.97	(0.92–1.02)	0.86–1.08	7
E6/E7 based PCR/cyto (ASC-US+)	CIN2+	1.63	(1.27–2.09)	-	0.90	(0.88–0.92)	-	1
APTIMA/cyto (ASC-US+)	CIN2+	1.35	(1.20–1.55)	1.33–1.50	0.98	(0.94–1.02)	0.96–1.00	2
Pretect HPV-Proofer	CIN2+	1.22	(0.79–1.87)	-	1.00	(0.97–1.04)	-	1

[a] When 2 values are given for number of studies: 1st value = number of studies where relative sensitivity was assessable and the 2nd = number of studies where the relative specificity could be assessed.
[b] The meta-analysis of relative sensitivity includes randomised controlled trials with a control arm where only cytology was used, the meta-analysis of relative specificity does not include these randomised controlled trials.
[c] Restricted to studies conducted in North America or Europe after exclusion of Kitchener HC *et al.* [14].
[d] Exclusion of Sankaranarayanan R *et al.* [8].
ASC-US: Atypical squamous cells of undetermined significance; HC2: Hybrid Capture® 2; LSIL: Low-grade squamous intraepithelial lesion.

The relative sensitivity and specificity of several HPV assays (GP5+/6+ PCR, careHPV™, cobas® 4800, Abbott RT PCR, PapilloCheck®, APTIMA®, Cervista®) compared to HC2 is displayed in Table 4. The 95% CIs around the accuracy measures always included unity at the exception of APTIMA, which was on average more specific for CIN2+ than HC2 (ratio of 1.07; 95% CI: 1.05–1.08).

3.3.1.3. Relative accuracy of screening with HPV testing, cytology and the combination of both tests estimated from baseline results in randomised trials. The characteristics and baseline results from eight randomised trials comparing accuracy for high-grade CIN of screening with cytology, hrHV testing or the combination of both have been published some years ago [7–14]. Fig. 5 summarises the relative sensitivity for CIN2+ (or detection rate ratio [DR ratio]) in the eight trials, of HPV- versus cytology-based screening in the trials, grouped by cytology method and by developing versus industrialised countries. HPV-based screening showed consistently and significantly higher sensitivity for CIN2+ than conventional cytology in the industrialised countries (ratio of 1.39; 95% CI: 1.23–1.57, p for inter-study heterogeneity = 0.57). No higher sensitivity was observed in the Indian trial [8] and in the ARTISTIC trial [14]. Possible reasons for the lack of higher sensitivity could be attributed to reference standard misclassification in the Indian trial, and over-diagnosis of lesions detected by liquid-based cytology and incomplete follow-up of HPV-positive women [15].

The pooled DR ratio, considering CIN3+ as outcome, was slightly lower than for CIN2+ (1.28; 95% CI: 1.09–1.51). In all of the European trials, hrHPV screening detected more CIN3+ than conventional cytology. However, this was significant only in the second phase of the Italian NTCC study (RR: 2.06; 95% CI: 1.16–3.68) [12] (Fig. 5). In all trials, with an experimental arm with combined screening, adding cytology to hrHPV testing yielded only a minor and statistically insignificant increase in sensitivity (pooled DR ratios of 1.06 [95% CI: 0.97–1.16] and 1.04 [95% CI: 0.92–1.17] for CIN2+ and CIN3+, respectively) (Fig. 6).

3.3.2. Longitudinal outcomes of RCTs

3.3.2.1. Reduction in CIN3+ and cancer observed in the second screening round. For four trials, results of the second screening round were published. Despite different policies of follow-up of screen-positive women, a remarkably consistent reduction of CIN3+ lesions was found among women who had an HPV-negative result versus those who had a negative cytology at enrollment (DR ratio of 0.43; 95% CI: 0.33–0.56) (Fig. 7) [11,14,17,18]. In addition, in three trials, the detection rate of invasive cancer at round 2 was significantly lower among women who were HPV-negative at initial screening (DR ratio of 0.13; 95% CI: 0.04–0.44) [11,17,18].

The Indian screening trial did not involve subsequent screening rounds but consisted in passive follow-up through routine cancer registration after a once-in-a-lifetime screening. A statistically

Table 4
Relative accuracy of other HPV tests compared to the HC2 assay (at RLU≥1) to find underlying CIN2 or CIN3 or worse in primary screening.

Comparison	Outcome	Relative sensitivity (95% CI)		Range	Relative specificity (95% CI)		Range	No. of studies [Reference]
GP5+/6+/HC2	CIN2+	1.00	(0.96-1.04)	0.98-1.00	0.99	(0.91-1.07)	0.95-1.1	2 [4,51]
careHPV (at RLU≥0.5)/HC2		0.93	(0.85-1.01)	-	0.98	(0.96-1.01)	-	1 [52]
Cobas-4800/HC2		0.98	(0.88-1.10)	-	1.00	(0.98-1.03)	-	1 [24]
Abbott RT PCR/HC2		1.00	(0.96-1.04)	0.99-1.03	1.01	(0.99-1.03)	1.00-1.02	1 [26,27]
Papillochek/HC2		0.98	(0.96-1.01)	-	0.99	(0.98-1.00)	-	1 [25]
APTIMA/HC2		1.02	(0.86-1.20)	0.95-1.12	1.07	(1.05-1.08)	1.06-1.08	2 [49,53]
careHPV (at RLU≥0.5)/HC2	CIN3+	0.91	(0.76-1.09)	-	0.98	(0.96-1.01)	-	1 [52]
Cervista/HC2		0.97	(0.93-1.02)	-	1.03	(1.02-1.04)	-	1 [54]
Abbott RT PCR/HC2		1.00	()[a]	-	1.02	(1.00-1.03)	-	1 [26]
MALDITOF/HC2		0.96	(0.92-1.01)	-	1.02	(1.01-1.03)	-	1 [54,55]
APTIMA/HC2		1.02	(0.93-1.11)	1.00-1.07	1.07	(1.05-1.08)	1.06-1.08	2 [49,53]

[a] 95% CI not computable since Abbott RT PCR and HC2 both showed 100% sensitivity for CIN3+.
CI: Confidence interval; CIN: Cervical intraepithelial neoplasia; HC2: Hybrid Capture® 2; RLU: Relative light unit.

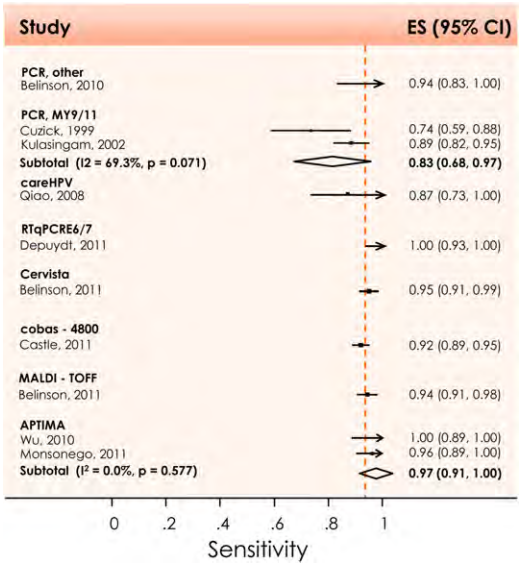


Fig. 4. Meta-analysis of the sensitivity for detecting CIN3+ in primary cervical cancer screening, using other tests than HC2, by test system. See web table liststudies.xls-screening for publication information. CI: Confidence interval; ES: Estimate of sensitivity; HC2: Hybrid Capture® 2; I^2: Percentage of total variation across studies due to heterogeneity; p: Test for inter-study heterogeneity.

significant reduction in incidence of advanced cervical cancer (stage II+) and of mortality from cervical cancer was observed in women in the HC2-screening arm compared to the control arm where no screening was offered (hazard ratios of 0.47 [95% CI: 0.35–0.63] and 0.52 [95% CI: 0.33–0.83], respectively) [19]. No significant reductions were observed in the study arms where cytology or VIA screening were applied.

3.3.2.2. Over-diagnosis of regressive lesions. Over-diagnosis can be assessed from differences between screening arms in the amount of high-grade CIN detected over successive screening rounds. The only study that applied HPV testing at the second screening round, POBASCAM, did not observe a difference between study arms in the detection of CIN2+ and CIN3+ over the first two rounds, but observed an excess of CIN2 in the HPV arm (ratio vs cytology group 1.33; 95% CI: 1.06–1.68) [20]. HPV testing was also used, however, in controls at the second round, allowing a more unbiased estimation of over-diagnosis. Several other European randomised controlled trials applied cytology at the second round. The Swedescreen trial found no difference concerning CIN3+ (ratio HPV vs. cytology group 1.04, $p = 0.20$) but some excess of CIN2 in the HPV arm (ratio 1.56, $p = 0.04$) [11]. In ARTISTIC, the ratio was 1.18 (95% CI: 0.90–1.55) for CIN2 and 0.85 (95% CI: 0.67–1.08) for CIN3+ [16]. NTCC found larger increases in the HPV group. Among women aged 35–60 years at recruitment, the HPV vs. cytology ratio was 1.65 (95% CI: 1.21–2.26) for CIN3 and 1.68 (95% CI: 1.25–2.26) for CIN2 [17]. The increase was larger among younger women (age 25–34 years at recruitment). The HPV vs. cytology ratio was 2.14 (95% CI: 1.28–3.59) for CIN3 in phase 2 of NTCC (the ratio was 0.99; 95% CI: 0.61–1.65 in phase 1, which did not show earlier detection of CIN3). For CIN2, the ratio was 3.11 (95% CI: 2.20–4.39) for both phases pooled. However, similar (statistically significant) values were observed in both phases, i.e., with direct referral and with cytologic triage.

The difference between the Swedish and Italian trials could be explained either by greater over-diagnosis with direct referral (second phase Italian trial) than with triage (triage by cytology + type-specific HPV persistence in Swedescreen), by larger lead time gain with direct referral than with cytologic triage or, most plausibly, by both. The large ratio (>3) observed for CIN2 below age 35 in the Italian study strongly suggests that HPV screening in younger women leads to serious over-diagnosis of regressive CIN2 [17].

3.3.3. Longitudinal outcomes of cohort studies

Fig. 8 shows the cumulative risk to develop CIN3+ in the next 5 years according to the initial combined HPV and cytology status (pooled from five European and two American screening cohorts) [21–23]. The risk is very low for HPV-negative women (0.2% and 1.2%, for women without or with cytologic abnormalities, respectively). However, for women who were hrHPV

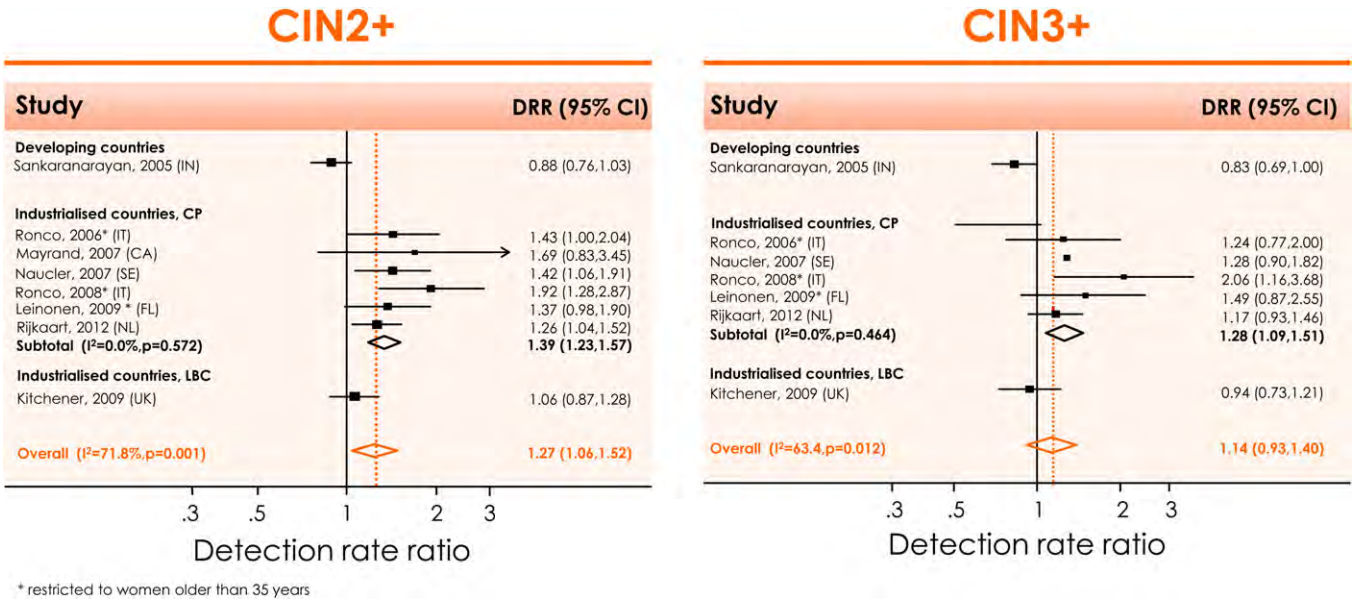


Fig. 5. Meta-analysis of randomised trials: DRR of CIN2+ (left) and CIN3+ (right) identified by hrHPV testing versus cytology in the first screening round. Studies are grouped by type of cytology (conventional cytology [CP] or liquid-based cytology [LBC]) and by industrialised versus developing countries. See web table liststudies.xls-screening for publication information. CIN: Cervical intraepithelial neoplasia; DRR: Detection rate ratio; I^2: the percentage of total variation across studies due to heterogeneity; p: test for inter-study heterogeneity.

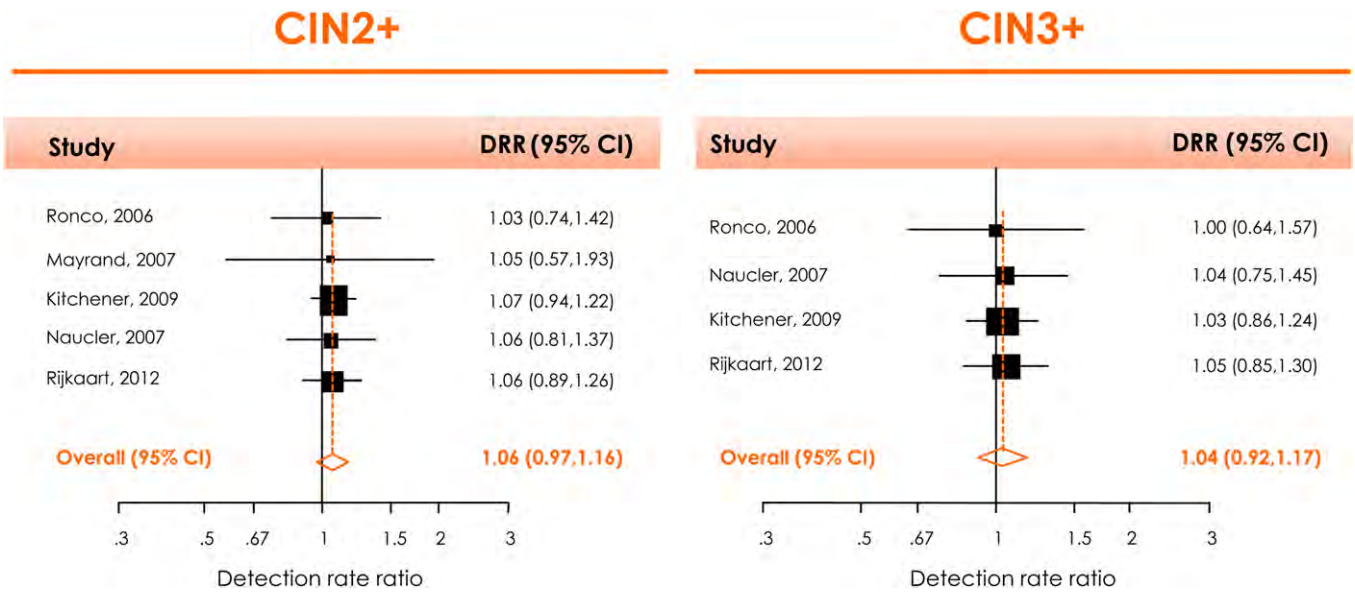


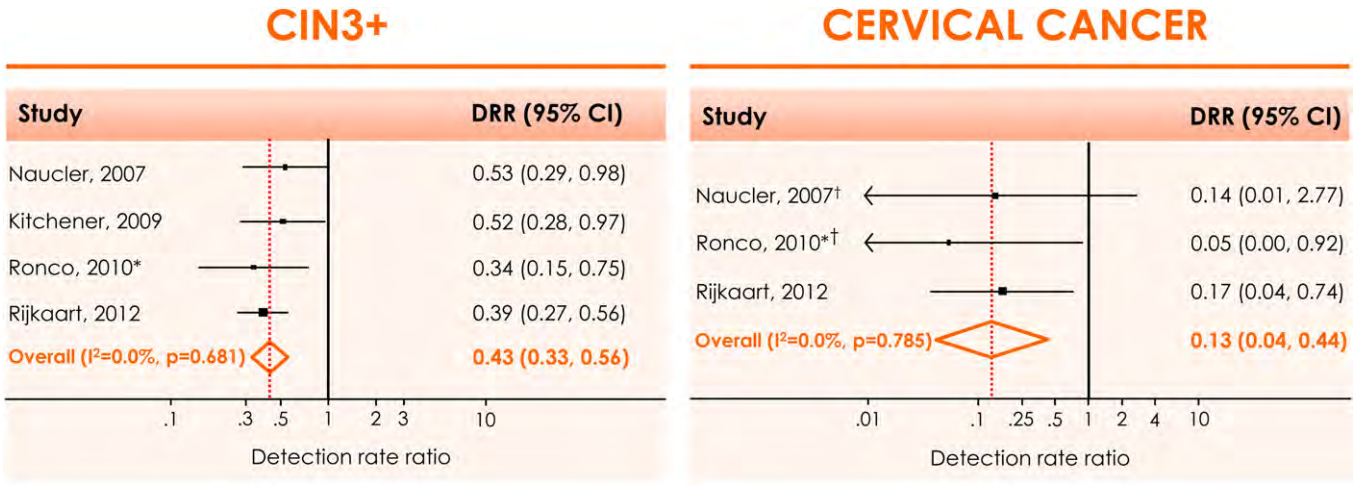
Fig. 6. Relative sensitivity of combined testing (HPV+cytology) vs. HPV testing only. See web table liststudies.xls-screening for publication information. CI: Confidence interval; CIN: Cervical intraepithelial neoplasia; DRR: Detection rate ratio.

positive, this risk was substantially higher: 6% or 17%, for women without or with ASC-US or worse cytology at baseline, respectively. The Portland study also provided outcomes regarding the risk of developing CIN3+ over 10 years in a population screened by cytology only, where hrHPV testing was performed at baseline: 10.2% (if ASC-US+), 6.9% (if hrHPV+), 6.8% (if hrHPV+ or ASC-US+), 1.4% (if cytology-negative), 0.9% (if hrHPV-negative) and 0.8% (if negative for both cytology and hrHPV) [23]. In the Kaiser-Permanente cohort, where women of 30 years or older were screened by cytology and HC2, the 5-year cumulative incidence of invasive cervical cancer was: 0.90% if positive for both hrHPV and cytology, 0.54% if hrHPV positive and cytology-negative, 0.16% if hrHPV-negative and cytology positive and 0.016% if negative for both [22].

3.3.4. Clinical validation for use in screening

Three hrHPV DNA tests (cobas® 4800, Abbott RT hrHPV PCR and Papillocheck®) have been evaluated clinically with the purpose of use in primary screening according the equivalence criteria of Meijer CJLM *et al.* [4,24–27]. All the three tests fulfilled minimal requirements regarding relative sensitivity or specificity compared to HC2 or GP5+/6+ PCR (Table 5). The reproducibility (hrHPV positive/negative) using cobas® 4800 and Abbott RT PCR largely exceeded the required cut-off but was not documented for Papillocheck®.

The topic of HPV screening in developing countries and HPV testing using samples collected by the woman is discussed by Cuzick J *et al.*, Vaccine, this issue [28].



* restricted to women of 35 years or older.
† continuity correction (+.5 in each cell because of zero cancer cases among HPV-negative women).

Fig. 7. Meta-analysis of the main outcomes from randomised trials comparing HPV- and cytology-based cervical cancer screening. Relative detection rate of CIN3+ (left panel) and cervical cancer (right panel), observed in the second screening round among women who were HPV-negative versus cytology-negative at enrolment. See web table liststudies.xls-screening for publication information. CI: Confidence interval; CIN: Cervical intraepithelial neoplasia; DRR: Detection rate ratio; I^2: the percentage of total variation across studies due to heterogeneity; p: test for inter-study heterogeneity.

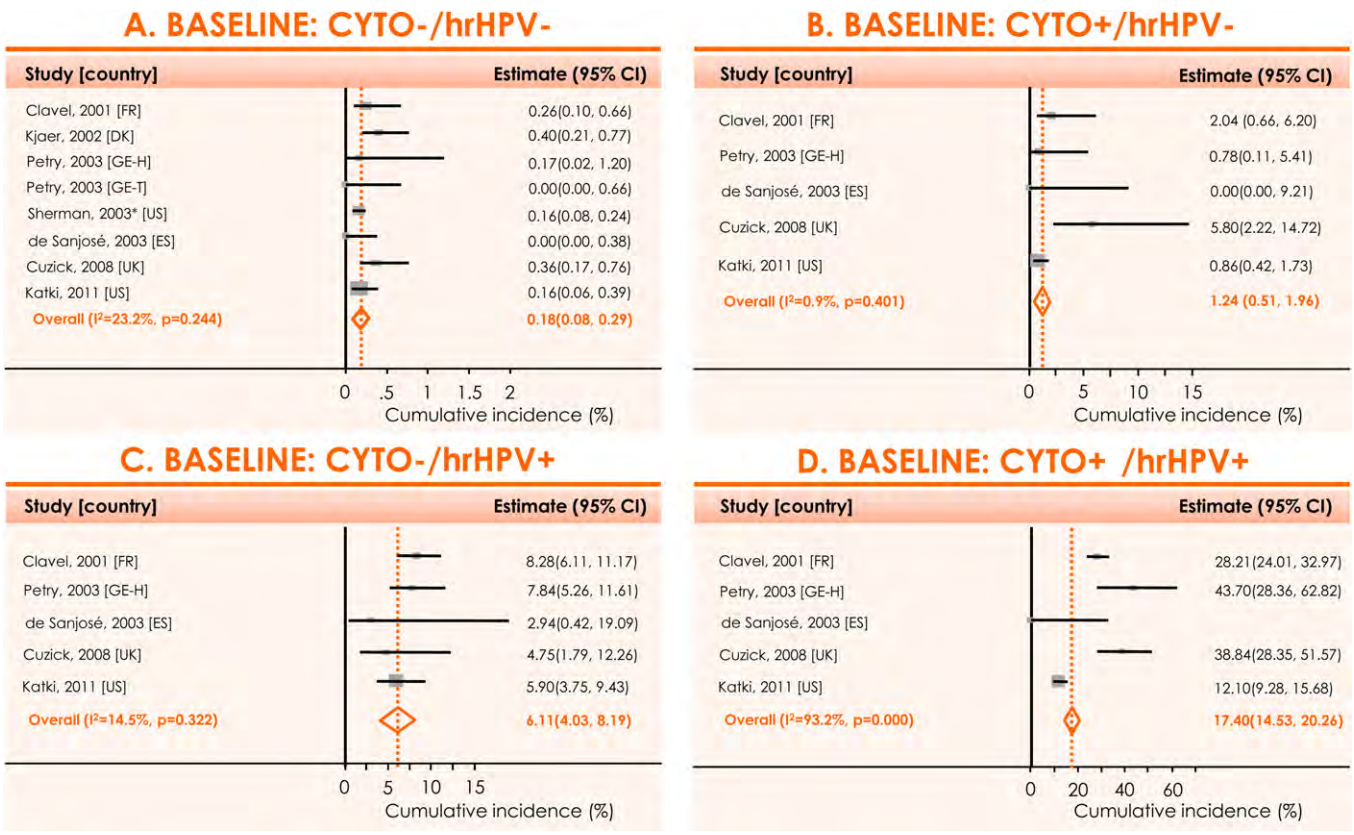


Fig. 8. Five-year cumulative incidence of CIN3+ among women attending cervical cancer screening who are at baseline (A) cyto-negative/hrHPV-negative; (B) cyto-positive/hrHPV-negative; (C) cyto-negative/hrHPV-positive; and (D) cyto-positive/hrHPV-positive. Pooled from 6 European and 2 American screening cohorts. * The Portland study reported cumuluative incidences over ~4 years (Sherman ME *et al.* [23]). See web table liststudies.xls-screening for publication information. CI: Confidence interval; CIN: Cervical intraepithelial neoplasia; hrHPV: High-risk human papillomavirus; I^2: the percentage of total variation across studies due to heterogeneity; p: test for inter-study heterogeneity.

4. Discussion

In this paper, a large body of data from clinical trials has been synthesised. Systematic assessment of the evidence base is an essential step for developing practice guidelines.

4.1. Triage of women with minor abnormal cytology

Testing with the HC2 assay has been shown to be an efficient triage tool for ASC-US cytology in the framework of cytology-based screening [29] and has been widely implemented in clinical practice. The updated meta-analysis corroborates the conclusions from previous reviews indicating better performance of HC2 to triage women with ASC-US (higher sensitivity, similar specificity) compared to repeat cytology [30]. New findings concern the accuracy of other tests, for which we propose the following thumb rule allowing their use in triage of equivocal cytology: a) clinical sensitivity of at least 90% for CIN2+ with relative sensitivity vs. HC2 not significantly lower than unity and a lower CI not lower than 0.90; b) relative specificity not significantly below unity with a lower CI not below 0.95. Abbott RT PCR, Papillocheck®, Cervista® and APTIMA® fulfil these criteria. However, it should be recognised that the specificity for excluding high-grade CIN2 (in the range 40–60% and consequently also the positive predictive value (range 20–30% for CIN2+, range 10–15% for CIN3+)) is not optimal, allowing for the use of more specific markers, such as genotyping for HPV16 and 18 or RNA testing for 5 types. Being positive for such markers allows identification of women at the highest risk (PPV for CIN2+ > 30%) warranting more intensive follow-up. However, women testing negative for these markers cannot be referred back to routine screening but require further follow-up. Interestingly, a recent meta-analysis showed that triage of ASC-US with p16INK4a immunocytochemistry was not less sensitive but was substantially more specific for identifying CIN2+ compared to HC2 (specificity ratio of 1.82 [95% CI: 1.57–2.12]) [31].

The findings on triage of women with LSIL are different from previous meta-analyses, which concluded that triage with HC2 was not more sensitive but substantially less specific than repeat cytology [2]. Based on those previous findings, HC2 was not recommended to manage women with low-grade cytologic lesions in European or American guidelines [32,33]. In the current review, significantly higher sensitivity of HC2 for CIN2+ and CIN3+ was demonstrated. This new finding could justify recommending the use of HC2 or other hrHPV DNA assays to decide whether women need referral to colposcopy. However, the high prevalence of hrHPV DNA among women with LSIL limits its utility for triage for this cytology category [29]. Therefore, the trade-off to use hrHPV tests in LSIL management should be based on local cost-effectiveness analyses, the local prevalence of HPV in LSIL and compliance of women with follow-up recommendations. In the USA, adult women with LSIL are referred directly to colposcopy [33]. If colposcopy and/or biopsy are normal or only reveal CIN1, an HPV test at 12 months or two repeat smears after the initial LSIL smear are recommended. Postponing triage is another option, allowing viral clearance, which over a period of 6 to 12 months can vary from 18–45% [34] and which reduces the need for colposcopy. hrHPV DNA testing certainly is not useful in young women (<25 years) with LSIL, given the very high prevalence of HPV infection and the very high probability of regression unrelated to initial HPV status [35]. HPV triage in older women, using a higher viral load cut-off, might be an acceptable management option, but more age-specific accuracy data are needed to define the exact age cut-off. In the

Table 5
Clinical validation of HPV DNA assays for use in primary cervical cancer screening: sensitivity and specificity of the assay to identify CIN2+ lesions compared to HC2 or GP5+/6+ PCR, reproducibility of hrHPV DNA detection (as proposed by Meijer CJLM *et al.* [4]).

	Evaluated	Absolute		Reference	Absolute		Relative		Non inferiority test		Reproducibility (95% CI)	
Study [reference]	assay	sensitivity	specificity	assay	sensitivity	specificity	sensitivity	specificity	p_sens	p_spec	Overall	Kappa
Meijer CJLM *et al.* [4]	GP5+/6+[a]	98.7%	96.0%	HC2	98.7%	94.1%	1.00	1.02	0.004	0.011		
Carozzi FM *et al.* [27]	Abbott RT hrHPV PCR	96.4%	92.3%	HC2	97.6%	92.6%	0.99	1.00	0.004	0.008	98.5% (97–99%)	0.97 (0.88–1.0)
Poljak M *et al.* [26]	Abbot RT hrHPV PCR	100.0%	93.3%	HC2	97.4%	91.8%	1.03	1.02	0.004	0.009	100.0% (99–100%)	1.00 (0.90–1.0)
Heideman DA *et al.* [24]	cobas-4800	90.0%	94.6%	HC2	91.7%	94.4%	0.98	1.00	0.005	0.007	98.3% (97–99%)	0.96 (0.87–1.0)
Hesselink AT *et al.* [25]	Papillocheck	95.8%	96.7%	GP5+/6+[a]	96.4%	97.7%	0.99	0.99	0.004	0.007	-	-

[a] PCR system amplifying HPV L1 DNA, including identification of 14 hrHPV types.
CI: Confidence interval; CIN: Cervical intraepithelial neoplasia; HC2: Hybrid Capture® 2; hrHPV: High-risk human papillomavirus; PCR: Polymerase chain reaction.

USA, LSIL triage with hrHPV testing is only recommended for postmenopausal women [36]. A recent meta-analysis revealed that the APTIMA® test showed similar sensitivity but higher specificity compared to HC2 [5]. HPV DNA genotyping for types 16 and 18, mRNA testing for 5 types, and p16INK4a and p16/Ki67 immunochemistry [31] show substantially higher specificity than hrHPV DNA assays but are less sensitive. These methods may be considered in a two-step management strategy, with referral and repeat testing in triage-positive and triage-negative women, respectively.

4.2. Follow-up after treatment of high-grade CIN

Women treated for cervical precancer must be followed carefully in order to monitor possible recurrence. The rate of residual or recurrent high-grade CIN, evaluated over 2 years or less, varied in the 15 included studies from 4–18%, with an average of 8%. The risk of recurrent CIN2+ is higher in women older than 50 years [37], which is consistent with the observation that viral persistence increases with age [38]. Recently, pooled long-term follow-up data indicate that women having been treated for CIN are still at increased risk for subsequent invasive cervical cancer compared to the general population, during at least 10 years and maybe up to 20 years after treatment [39]. Finding an indicator that predicts successful outcome, allowing shortening of the follow-up period, would be particularly helpful. Currently available data suggest that HPV testing picks up residual disease quicker and with higher sensitivity and similar specificity compared to follow-up cytology or the histological assessment of the section margins. Only one study addressed the predictive value of consecutive testing for an outcome over 10 years [40]. One negative test at 6 months after treatment resulted in a 10-year risk of CIN3+ of 2.1% (if hrHPV-negative), 2.8% (if cytology-negative) and 1.4% (if combined hrHPV- and cytology-negative). Three consecutive negative cytology results (6, 12 and 24 months) or two negative co-tests (HPV-negative and cytology-negative, at 6 months and 12 months) were associated with a low risk of residual CIN3+ of 0.7% and 0.0%, respectively, which was not different from the general population. On the other side, one positive test at 6 months increased the risk of CIN3+ dramatically: 29%, 13% and 23%, if hrHPV-positive, cytology ASC-US+, or one of both tests positive, respectively [40].

4.3. Primary screening for cervical (pre-)cancer

The updated meta-analyses confirm that hrHPV testing is substantially more sensitive than cytology in identifying underlying CIN2+ and CIN3+. However, one drawback is the lower specificity, in particular among younger women where HPV infection usually is transient. A wide spread of accuracy estimates is observed in certain developing countries, possibly explainable by variability in the reference standard. Further variation is observed according to the intensity of verification and methods to adjust for verification bias [2,41].

The higher cross-sectional sensitivity of hrHPV testing for detecting CIN2+ and CIN3+ provides insufficient evidence that HPV-based screening will decrease the incidence of cervical cancer more than cytology-based screening. Most CIN2 and CIN3 lesions clear and it cannot be excluded from cross-sectional studies that HPV tests just pick up more regressive disease. For, this reason, hrHPV-based cervical cancer screening has not been recommended yet in the 2nd edition of the European Guidelines for Cervical Cancer Screening, considering evidence available in 2006 [42]. A significantly lower cumulative incidence of CIN3+, observed in the second screening round in women who were hrHPV-negative versus cytology-negative, at enrolment, was an anticipated condition to propose switching to HPV-based screening. Today, this type

of evidence is available. In spite of the differences in design and management of screen-positive women, a significant reduction in the risk of CIN3+ and even of cancer was consistently demonstrated in the randomised trials that published second round outcomes, (see Fig. 7).

A low cumulative incidence of CIN3+ and cancer associated with a negative hrHPV DNA test was also observed in screening cohorts (Fig. 8). This risk was only slightly lower in women being co-test negative compared to those were hrHPV-negative but cytology-positive, suggesting that screening with only a hrHPV test without cytology offers similar protection. However, in hrHPV positive women, cytology clearly stratifies the risk for CIN3+ and cancer in a very high-risk group (17% CIN3+ and 1% cancer, over 5 years) and a moderately high-risk group (6% and 0.5%, respectively), underscoring the potential of cytology as an effective triage method in an HPV-based screening strategy [20,22,43]. Being hrHPV-positive was more predictive for CIN3+ over a 5-year period than ASC-US+ cytology in the Kaiser-Permanente cohort [22]. This finding may look peculiar, given the lower cross-sectional specificity of hrHPV testing but can be explained by the development of more new CIN3 lesions over time in hrHPV-positive compared to cytology-positive women [22].

Besides cytology, genotyping for HPV type 16 or 16 and 18, combinations of HPV16/18 genotyping and cytology, or HPV retesting to detect persistent infections, also stratify the risk for CIN3+ and can be used in triage of hrHPV-positive women [20,44–46]. The 10-year cumulative risk associated with HPV16, HPV18 or other high-risk HPV infection, among women included in the Portland study was 17%, 14% and 3%, respectively [47]. In the ATHENA trial, identification of HPV16 or HPV18 with the Cobas® 4800 test showed similar sensitivity and PPV for CIN3+ (60% and 16%, respectively) as cytology at ASC-US+ cut-off (53% and 14%, respectively) to triage hrHPV-positive women [48]. Other candidate markers for triage, which could be considered, but for which evidence is currently still insufficient, are: testing for viral RNA of a limited number of types, p16INK4a immunocytochemistry or p16INK4a i67 double staining. Currently, no marker, or combination of markers, allow perfect reflex triage. hrHPV-positive women being cytology-negative or HPV16/18-negative must be kept under surveillance and invited for repeat testing. Additionally, HPV screening using a more specific test, such as the APTIMA RNA assay [49] or HC2 at a higher viral load cut-off [50] increases PPV with no or a small losses in cross-sectional sensitivity.

HC2, GP5+/6+, Cobas® 4800 and the Abbott RT hrHPV PCR can be considered as clinically validated assays usable for primary HPV-based cervical cancer screening. As indicated in Tables 1 and 3, and in Fig. 4, several other hrHPV DNA tests may be sufficiently accurate but are not yet validated using the criteria of Meijer CJLM *et al.* [4]. Besides accuracy, other characteristics will also have to be taken in account in future choices of tests, such as high-throughput capacity, costs, applicability on samples taken by the woman, and the possibility to perform ancillary triage tests. The equivalence criteria based on relative cross-sectional accuracy [4] are not valid for other molecular markers (for instance over-expressed proteins or viral RNA). These markers should reach the same level of evidence as required for the original HPV DNA screening tests: the demonstration of lower cumulative incidence of CIN3+ after a baseline negative screening test result versus after a negative cytology or a similar cumulative incidence of CIN3+ compared to that seen after a negative result with a clinically validated hrHPV DNA test.

5. Conclusions

The evidence base supports the notion that hrHPV tests can provide an advantage in triage of women with equivocal cytologic lesions and in the follow-up after treatment of cervical precancer. However, in triage of LSIL, more specific tests are preferred. HPV DNA or RNA genotyping for a limited number of HPV types allow identification of women at highest risk, but women testing negative for these tests still will need further testing.

The evidence base indicates that hrHPV DNA testing with a clinically validated assay is more effective than cervical cytology in primary screening of women of 30 years and older and that screening intervals for HPV-negative women can be safely extended to at least 5 years. Reflex cytology and genotyping for HPV16 or HPV18 can be used to triage hrHPV-positive women. The possible advantages offered by HPV-based screening require a well organised programme with good compliance with screening and triage policies.

Acknowledgements

We acknowledge and thank the following authors who have provided additional non-published data which were used in the meta-analyses: S. Andersson, S. Belinson, J. Cage, F. Coutlée, K. Cuschieri, J. Cuzick, M. Fröberg, M. Guo, P. Halfon, S. Hibbits, M.Z. Huang, S. Huang, S. Kulasingam, A. Lytwyn, J. Monsonego, M. Nasioutziki, P. Naucler, A. Pista, R. Pretorius, R. Sankaranarayanan, S. Ratnam, M. Reuschenbach, G. Ronco, M. Stoler, K. Syrjanen, A. Szarewski and N. Wentzensen. L. Houthuys is acknowledged for bibliographic support.

MA received financial support from: (1) Directorate of SANCO of the European Commission, Luxembourg, Grand-Duchy of Luxembourg), through the ECCG project (European Cooperation on development and implementation of Cancer screening and prevention Guidelines, grant no. 2006322), coordinated by IARC, Lyon, France; (2) the 7th Framework Programme of DG Research of the European Commission through the PREHDICT project (grant No. 242061, coordinated by the Vrije Universiteit Amsterdam, the Netherlands); (3) the Belgian Foundation Against Cancer (Brussels, Belgium); (4). Gynaecological Cancer Cochrane Review Collaboration (Bath, UK).

Disclosed potential conflicts of interest

MA: See acknowledgements.

GR, AA, GO, GK, RS, JP: Have disclosed no potential conflicts of interest.

CJLMM: Scientific advisory board (Qiagen), speakers office (GSK, Roche), research grant (Abbott), shares minority (Self-Screen B.V)

MP: Advisory Board (Roche); Consultant (Abbott); Research Grants (Abbott, Roche); Speakers Bureau (Abbott, Roche); Travel Grants (Abbott, Roche).

PN: Has received lecture fees from Abbott Laboratories.

References

[1] Bosch FX, Lorincz AT, Munoz N, Meijer CJ, Shah KV. The causal relation between human papillomavirus and cervical cancer. J Clin Pathol 2002;55:244–65.
[2] Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. Vaccine 2006;24(Suppl 3):S78–89.
[3] Koliopoulos G, Arbyn M, Martin-Hirsch P, Kyrgiou M, Prendiville W, Paraskevaidis E. Diagnostic accuracy of human papillomavirus testing in primary cervical screening: a systematic review and meta-analysis of non randomised studies. Gynecol Oncol 2007;104:232–46.
[4] Meijer CJLM, Berkhof J, Castle PE, Hesselink AT, Franco EL, Ronco G, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. Int J Cancer 2009;124:516–20.
[5] Arbyn M, Roelens J, Cuschieri K, Cuzick J, Szarewski A, Ratnam S, et al. The APTIMA HPV assay versus the Hybrid Capture II test in triage of women with ASC-US or LSIL cervical cytology: a meta-analysis of the diagnostic accuracy. Int J Cancer 2012 May 18, http://dx.doi.org/10.1002/ijc.27636 [Epub ahead of print].

[6] Poljak M, Cuzick J, Kocjan BJ, Iftner T, Dillner J, Arbyn M. Nucleic acid tests for the detection of alpha human papillomaviruses. Vaccine 2012;30(S5):F100–6.
[7] Davies P, Arbyn M, Dillner J, Kitchener HC, Ronco G, Hakama M. A report on the current status of European research on the use of human papillomavirus testing for primary cervical cancer screening. Int J Cancer 2006;118:791–6.
[8] Sankaranarayanan R, Nene BM, Dinshaw KA, Mahe C, Jayant K, Shastri SS, et al. A cluster randomized controlled trial of visual, cytology and human papillomavirus screening for cancer of the cervix in rural India. Int J Cancer 2005;116:617–23.
[9] Ronco G, Segnan N, Giorgi-Rossi P, Zappa M, Casadei GP, Carozzi F, et al. Human Papillomavirus testing and liquid-based cytology in primary cervical screening: results at recruitment from the New Technologies for Cervical Cancer randomized controlled trial. J Natl Cancer Inst 2006;98:765–74.
[10] Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. N Engl J Med 2007;357:1579–88.
[11] Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgren K, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. N Engl J Med 2007;357:1589–97.
[12] Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, 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;100:492–501.
[13] Leinonen M, Nieminen P, Kotaniemi-Talonen L, Malila N, Tarkkanen J, Laurila P, et al. Age-specific evaluation of primary human papillomavirus screening vs conventional cytology in a randomized setting. J Natl Cancer Inst 2009;101:1612–23.
[14] Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. Lancet Oncol 2009;10:672–82.
[15] Arbyn M, Ronco G, Meijer CJLM, Naucler P. Trials comparing cytology with human papillomavirus screening. Lancet Oncol 2009;10:935–6.
[16] Bulkmans NW, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJ, Bulk S, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. Lancet 2007;370:796–802.
[17] Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. Lancet Oncol 2010;11:249–57.
[18] Rijkaart DC, Berkhof J, Rozendaal L, van Kemenade FJ, Bulkmans NW, Heideman DA, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. Lancet Oncol 2012;13:78–88.
[19] Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. HPV screening for cervical cancer in Rural India. N Engl J Med 2009;360:1385–94.
[20] Rijkaart DC, Berkhof J, van Kemenade FJ, Coupe VM, Hesselink AT, Rozendaal L, et al. Evaluation of 14 triage strategies for HPV DNA-positive women in population-based cervical screening. Int J Cancer 2012;130:602–10.
[21] Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. BMJ 2008;337:a1754.
[22] Katki HA, Kinney WK, Fetterman B, Lorey T, Poitras NE, Cheung L, et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. Lancet Oncol 2011;12:666–72.
[23] Sherman ME, Lorincz AT, Scott DR, Wacholder S, Castle PE, Glass AG, et al. Baseline cytology, human papillomavirus testing, and risk for cervical neoplasia: A 10-Year cohort analysis. J Natl Cancer Inst 2003;95:46–52.
[24] Heideman DA, Hesselink AT, Berkhof J, van Kemenade F, Melchers WJ, Daalmeijer NF, et al. Clinical validation of the cobas(R)4800 HPV Test for cervical screening purposes. J Clin Microbiol 2011;49:3983–5.
[25] Hesselink AT, Heideman DA, Berkhof J, Topal F, Pol RP, Meijer CJ, et al. Comparison of the clinical performance of PapilloCheck human papillomavirus detection with that of the GP5+/6+-PCR-enzyme immunoassay in population-based cervical screening. J Clin Microbiol 2010;48:797–801.
[26] Poljak M, Ostrbenk A, Seme K, Ucakar V, Hillemanns P, Bokal EV, et al. Comparison of clinical and analytical performance of the Abbott RealTime High Risk HPV test to the performance of Hybrid Capture 2 in population-based cervical cancer screening. J Clin Microbiol 2011;49:1721–9.
[27] Carozzi FM, Burroni E, Bisanzi S, Puliti D, Confortini M, Giorgi Rossi P, et al. Comparison of clinical performance of Abbott RealTime High Risk HPV test with that of Hybrid Capture 2 assay in a screening setting. J Clin Microbiol 2011;49:1446–51.
[28] Cuzick J, Bergeron C, von Knebel-Doeberitz M, Patti Gravitt, Jose Jeronimo, Attila T, Lorincz, et al. New technologies and procedures for cervical cancer screening. Vaccine 2012;30(S5):F107–16.
[29] Arbyn M, Martin-Hirsch P, Buntinx F, Van Ranst M, Paraskevaidis E, Dillner J. Triage of women with equivocal or low-grade cervical cytology results. A meta-analysis of the HPV test positivity rate. J Cell Mol Med 2009;13:648–59.
[30] Arbyn M, Buntinx F, Van Ranst M, Paraskevaidis E, Martin-Hirsch P, Dillner J. Virologic versus cytologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. J Natl Cancer Inst 2004;96:280–93.
[31] Roelens J, Reuschenbach M, von Knebel-Doeberitz M, Wentzensen N, Bergeron C, Arbyn M. p16INK4a immunocytochemistry versus HPV testing for triage of women with minor cytological abnormalities: A systematic review and meta-analysis. Cancer 2012, prepub-ahead of print.
[32] Jordan J, Arbyn M, Martin-Hirsch P, Schenck U, Baldauf JJ, Da Silva D, et al. European guidelines for quality assurance in cervical cancer screening: recommendations for clinical management of abnormal cervical cytology, part 1. Cytopathology 2008;19:342–54.
[33] Wright Jr TC, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. Am J Obstet Gynecol 2007;197:346–55.
[34] Bais AG, Rebolj M, Snijders PJ, de Schipper FA, van der Meulen DA, Verheijen RH, et al. Triage using HPV-testing in persistent borderline and mildly dyskaryotic smears: Proposal for new guidelines. Int J Cancer 2005;116:122–9.
[35] Moscicki AB, Shiboski S, Hills NK, Powell KJ, Jay N, Hanson EN, et al. Regression of low-grade squamous intra-epithelial lesions in young women. Lancet 2004;364:1678–83.
[36] Solomon D, Papillo JL, Davey DD, Statement on HPV, test DNA, Am F utilization. Clin Pathol 2009;131:768–9.
[37] Flannelly G, Bolger B, Fawzi H, De Lopes AB, Monaghan JM. Follow up after LLETZ: could schedules be modified according to risk of recurrence? BJOG 2001;108:1025–30.
[38] Castle PE, Schiffman M, Herrero R, Hildesheim A, Rodriguez AC, Bratti MC, et al. A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste. Costa Rica J Infect Dis 2005;191:1808–16.
[39] Soutter WP, Sasieni P, Panoskaltsis T. Long-term risk of invasive cervical cancer after treatment of squamous cervical intraepithelial neoplasia. Int J Cancer 2005;118:2048–55.
[40] Kocken M, Helmerhorst TJ, Berkhof J, Louwers JA, Nobbenhuis MA, Bais AG, et al. Risk of recurrent high-grade cervical intraepithelial neoplasia after successful treatment: a long-term multi-cohort study. Lancet Oncol 2011;12:441–50.
[41] Cuzick J, Arbyn M, Sankaranarayanan R, Tsu V, Ronco G, Mayrand MH, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. Vaccine 2008;26(Suppl 10):K29–41.
[42] European Commission. European Guidelines for Quality Assurance in Cervical Cancer Screening. 2nd ed Luxembourg: Office for Official Publications of the European Communities; 2008.
[43] Rijkaart DC, Coupe VM, van Kemenade FJ, Heideman DA, Hesselink AT, Verweij W, et al. Comparison of Hybrid capture 2 testing at different thresholds with cytology as primary cervical screening test. Br J Cancer 2010;103:839–946.
[44] Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al. American Cancer Society American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin 2012.
[45] Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgren K, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. J Natl Cancer Inst 2009;101:88–98.
[46] Castle PE, Rodriguez AC, Burk RD, Herrero R, Wacholder S, Alfaro M, et al. Short term persistence of human papillomavirus and risk of cervical precancer and cancer: population based cohort study. BMJ 2009;339:b2569–84.
[47] Khan MJ, Castle PE, Lorincz AT, Wacholder S, Sherman M, Scott DR, et al. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. J Natl Cancer Inst 2005;97:1072–9.
[48] Castle PE, Stoler MH, Wright Jr TC, Sharma A, Wright TL, Behrens CM. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. Lancet Oncol 2011;12:880–90.
[49] Monsonego J, Hudgens MG, Zerat L, Zerat JC, Syrjänen K, Halfon P, et al. Evaluation of oncogenic human papillomavirus RNA and DNA tests with liquid based cytology in primary cervical cancer screening (The FASE study). Int J Cancer 2011;129:691–701.
[50] Rebolj M, Bonde J, Njor SH, Lynge E. Human papillomavirus testing in primary cervical screening and the cut-off level for hybrid capture 2 tests: systematic review. BMJ 2011;342:d2757.
[51] Kulmala SM, Syrjänen SM, Shabalova I, Petrovichev N, Kozachenko V, Podistov J, et al. Human papillomavirus testing with the hybrid capture 2 assay and PCR as screening tools. J Clin Microbiol 2004;42:2470–5.
[52] Qiao YL, Sellors JW, Eder PS, Bao YP, Lim JM, Zhao FH, et al. A new HPV-DNA test for cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. Lancet Oncol 2008;9:929–36.
[53] Wu R, Belinson SE, Du H, Na W, Qu X, Wu R, et al. Human papillomavirus messenger RNA assay for cervical cancer screening: the Shenzhen Cervical Cancer Screening Trial I. Int J Gynecol Cancer 2010;20:1411–4.
[54] Belinson JL, Wu R, Belinson SE, Qu X, Yang B, Du H, et al. A population-based clinical trial comparing endocervical high-risk HPV testing using hybrid capture 2 and Cervista from the SHENCCAST II Study. Am J Clin Pathol 2011;135:790–5.
[55] Belinson JL, Du H, Yang B, Wu R, Belinson SE, Qu X, et al. Improved sensitivity of vaginal self-collection and high-risk human papillomavirus testing. Int J Cancer 2012;130:1855–60.

IJC
International Journal of Cancer

The APTIMA HPV assay *versus* the Hybrid Capture 2 test in triage of women with ASC-US or LSIL cervical cytology: A meta-analysis of the diagnostic accuracy

Marc Arbyn[1], Jolien Roelens[1], Kate Cuschieri[2], Jack Cuzick[3], Ann Szarewski[3], Sam Ratnam[4], Miriam Reuschenbach[5], Suzanne Belinson[6], Jerome L. Belinson[6] and Joseph Monsonego[7]

[1] Unit of Cancer Epidemiology, Scientific Institute of Public Health, Brussels, Belgium
[2] Scottish HPV Reference Laboratory, Royal Infirmary of Edinburgh, Edinburgh, Scotland, United Kingdom
[3] Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, United Kingdom
[4] Public Health Laboratory, Department of Health, Division of Community Medicine, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
[5] Department of Applied Tumor Biology, Institute of Pathology, University of Heidelberg, Heidelberg, Germany
[6] Preventive Oncology International, Inc., Cleveland, OH, USA
[7] Institute of the Cervix, Paris, France

Testing for DNA of 13 high-risk HPV types with the Hybrid Capture 2 (HC2) test has consistently been shown to perform better in triage of women with cervical cytology results showing atypical squamous cells of undetermined significance (ASC-US) but often not in triage of low-grade squamous intraepithelial lesions (LSIL) detected in cervical cancer screening. In a meta-analysis, we compared the accuracy of the APTIMA HPV test, which identifies RNA of 14 high-risk HPV types, to HC2 for the triage of women with ASC-US or LSIL. Literature search-targeted studies where the accuracy of APTIMA HPV and HC2 for detection of underlying CIN2/3+ was assessed concomitantly including verification of all cases of ASC-US and LSIL. HSROC (Hierarchical Summary ROC) curve regression was used to compute the pooled absolute and relative sensitivity and specificity. Eight studies, comprising 1,839 ASC-US and 1,887 LSIL cases, were retrieved. The pooled sensitivity and specificity of APTIMA to triage ASC-US to detect underlying CIN3 or worse was 96.2% (95% CI = 91.7–98.3%) and 54.9%

Key words: cervical cancer screening, HPV, hybrid capture, APTIMA, triage, atypical squamous cells of undetermined significance, low-grade squamous intraepithelial lesions, diagnostic test accuracy, meta-analysis
Grant sponsors: European Commission, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines (ECCG), Belgian Foundation Against Cancer (Brussels, Belgium), Gynaecological Cancer Cochrane Review Collaboration (Bath, United Kingdom)
DOI: 10.1002/ijc.27636
History: Received 17 Feb 2012; Accepted 8 May 2012; Online 18 May 2012
Correspondence to: Marc Arbyn, Unit of Cancer Epidemiology, J. Wytsmanstreet 14, B1050 Brussels, Belgium, Tel.: +32-2-642-50-21, Fax: +32-2-642-54-10; E-mail: marc.arbyn@wiv-isp.be

Conflict of Interest M. Arbyn and J. Roelens were supported by (*i*) The Belgian Foundation Against Cancer, Brussels, Belgium; (*ii*) the European Commission through the PREHDICT Network (grant No. 242061), coordinated by the Free University of Amsterdam (The Netherlands), funded by the 7th Framework program of DG Research (Brussels, Belgium) and through the ECCG (grant no. 2006322, European Cooperation on development and implementation of Cancer screening and prevention Guidelines, via IARC, Lyon, France), funded by Directorate of SANCO (Luxembourg, Grand-Duchy of Luxembourg) and (*iii*) the Gynaecological Cancer Cochrane Review Collaboration (Bath, United Kingdom).
K. Cuschieri has been on the advisory board of Gen-Probe and has received project funding from Gen-Probe. The same applies to other companies including Qiagen, Roche and Abbott.
J. Cuzick is an occasional consultant for Gen-Probe, and his institution has research funding from Gen-Probe. The same applies to other HPV diagnostic companies: Abbott, BD, Qiagen and Roche.
The study of Wu (*Int J Gynecol Cancer*, 2010) received support from Qiagen and Gen-Probe. S. Belinson has no further conflict of interest to declare.
J.R. Belinson has received support (reagents, testing, funds or direct support and research) under the auspices of Preventive Oncology International Inc. from Hologic Inc., Qiagen, Gen-Probe Inc. and BGI Shenzhen.
A. Szarewski has no conflict of interest to declare.
The study of Ratnam (*Cancer Epidemiol Biomarkers Prev*, 2011) received support from Merck Frosst Canada Ltd. S. Ratnam has no further conflict of interest to declare.
M. Reuschenbach has no conflict of interest to declare.
J. Monsonego has received research funding from Gen-Probe Inc. and has participated at Advisory Board of Sanofi Pasteur MSD, Gen-Probe, Qiagen and Roche Diagnostics.

(95% CI = 43.5–65.9%), respectively. APTIMA and HC2 showed similar pooled sensitivity; however, the specificity of the former was significantly higher (ratio: 1.19; 95% CI = 1.08–1.31 for CIN2+). The pooled sensitivity and specificity of APTIMA to triage LSIL were 96.7% (95% CI = 91.4–98.9%) and 38.7% (95% CI = 30.5–47.6%) for CIN3+. APTIMA was as sensitive as HC2 but more specific (ratio: 1.35; 95% CI = 1.11–1.66). Results were similar for detection of CIN2 or worse. In both triage of ASC-US and LSIL, APTIMA is as sensitive but more specific than HC2 for detecting cervical precancer.

There is wide consensus that women with high-grade cytological lesions detected at cervical cancer screening need further diagnostic work-up.[1–4] However, follow-up guidelines in case of minor cytological abnormalities vary between countries and include cytological surveillance, triage using HPV DNA or RNA testing or referral for colposcopy.[1,5,6] Therefore, intensive research has been conducted over recent years to find improved management procedures.[7–9]

Women with minor cervical cytological lesions such as atypical squamous cells of undetermined significance (ASC-US) or low-grade squamous intraepithelial lesions (LSIL) have a small but significantly increased risk of developing cervical cancer compared to women with normal smears. Reviews of the natural history of cervical epithelial atypia or low-grade squamous lesions indicate that the 2-year cumulative risk of invasive cervical cancer is in the range of 0.10–0.25%.[10,11] Therefore, careful management of these lesions is warranted and the challenge is to discriminate accurately between those having underlying or incipient cancer precursors and those that can be released safely to routine screening 3 or 5 years later.

The Hybrid Capture 2 assay (HC2; Qiagen, GmbH, Hilden, Germany), which detects DNA of 13 high-risk types,[12] is the most frequently used HPV test worldwide. Systematic reviews[13–15] and a large randomized trial[16] have consistently shown superior accuracy of HC2 in triage of women with ASC-US cytology to detect underlying high-grade cervical intraepithelial neoplasia or worse (CIN2/3+) compared to repeat cytology. However, LSIL triage with HC2 did not show enhanced sensitivity and was substantially less specific than a repeat Pap test.[17,18] Restricting HPV-based triage of women with LSIL to older age groups has been proposed where specificity is higher. However, no universal age threshold can be identified.[19] In Italy, for instance, a rather good specificity was observed for LSIL triage with HC2 for women aged 35–60 years (61%) *versus* only 31% in the age group 25–34 years.[20] In contrary, in a large database in the United States, the specificity of HC2 in LSIL triage did not exceed 30% in women older than 50 years.[21]

In the current meta-analysis, we compared the accuracy of a new assay, the APTIMA HPV test (Gen-Probe Incorp, San Diego, CA), which identifies mRNA of 14 high-risk HPV types,[22] to HC2 for the triage of women with ASC-US or LSIL cytological findings.

Material and Methods

Clinical question

The purpose of the systematic review was to compare the accuracy of the HPV APTIMA assay and HC2 test to detect underlying cervical intraepithelial neoplasia grade 2 or worse (CIN2+) or CIN3+ in women with a cervical cytology specimen showing ASC-US or LSIL.

Literature search and eligibility criteria

PUBMED-Medline, EMBASE and CENTRAL were used for retrieving references. The search strategy used in PUBMED is provided in the Appendix. Furthermore, the abstract book from the 26th and 27th International Conferences of the Papillomavirus Society (IPV26, Montreal, 2010; IPV 27, Berlin, 2011) and the reference lists of retrieved reports were explored as well. The search started from 2007, the year of the first report on the APTIMA HPV test and included references published up to December 3, 2011. There was no language restriction. In addition, websites of the Food and Drugs Association (US FDA) and the European Medicine Agency were consulted to verify the presence of accuracy data of APTIMA submitted for authorization.

Two cytological groups were distinguished: (*i*) the ASC-US group, comprising ASCUS/ASC-US (defined according to the 1988/2001 editions of the Bethesda System[23,24]) or borderline dyskaryosis [defined according to the British Society of Clinical Cytology (BSCC) reporting guidelines] and (*ii*) the LSIL group, including LSIL[24] and mild dyskaryosis.[25] Authors were contacted to provide data separately for ASC-US or LSIL if the number of true and false positives and negatives could not be derived from the reports. The BSCC cytological categories[25] borderline and mild dyskaryosis were considered as equivalent to ASC-US and LSIL, respectively. Atypical squamous cells where high-grade lesions cannot be excluded and atypical glandular cells were not targeted in the review.

Evaluated tests

The evaluated index test was the APTIMA assay, which detects mRNA coding for the E6 and E7 viral proteins of HPV types 16, 18, 31, 33, 35, 38, 39, 45, 51, 52, 56, 58, 59, 66 and 68, accepting the cutoff for a positive test as proposed by the manufacturer.[26] The comparator test was the B-probe of HC2 test, which targets DNA sequences of the same HPV types with the exception of HPV 66 using the standard cutoff as positivity criterion (signal strength, relative light units $\geq$ 1, expressing semiquantitatively the viral load compared to a control sample with 1 pg of HPV DNA per milliliter).[12]

Reference standard

All women had to be submitted to verification with colposcopy, colposcopy-directed biopsies, possibly supplemented

with endocervical curettage when glandular disease was suspected. The histological interpretation of biopsies was considered as the outcome, accepting negative colposcopy as sufficient ascertainment for the absence of disease, when no biopsies were taken in the absence of colposcopic changes. Two levels of disease outcome were considered: CIN2+ and CIN3+. Adenocarcinoma *in situ* was included in the CIN3+ outcome.

Study selection, data extraction and statistical analysis

Two authors (M.A. and J.R.) independently checked eligibility of references and extracted the numbers of the true and false positives and negatives for each test, triage group and outcome. It was also noted whether tested women were attending routine screening or were referred to colposcopy for further assessment. The quality of the selected studies was evaluated using the QUADAS checklist.[27]

Accuracy parameters were pooled separately using a random effect model.[28] The absolute sensitivity and specificity of the tests were estimated jointly using *metandi*, a procedure in STATA, based on a two-level mixed logistic regression model, with independent binomial distributions for the true positives and true negatives conditional on the sensitivity and specificity in each study, and a bivariate normal model for the logit transforms of sensitivity and specificity between studies.[29,30] The relative sensitivity and specificity of APTIMA compared to HC2 were computed using *metadas*, a SAS macro for meta-analysis of diagnostic accuracy studies that allows the inclusion of test as a covariate making comparison of tests possible.[31,32] The influence of study characteristics, including QUADAS items and the used cytological classification system, on test accuracy were assessed using the HSROC regression of *metadas*.

Results

Selection of studies

The search in Medline yielded 225 articles, from which six relevant articles, evaluating the accuracy of both the APTIMA and HC2, could be identified.[26,33–37] One additional reference could be retrieved from the IPV26 abstract book,[38] and another was identified from the FDA web site http://www.accessdata.fda.gov/.[39] No additional references were identified from EMBASE or CENTRAL. Two articles that addressed only virological issues,[22,40] an article evaluating a prototype test different from the currently used APTIMA assay[41] and two secondary publications containing no new data[42,43] were excluded. In total, 1,839 women with ASC-US and 1,887 women with LSIL were enrolled in eight studies.[26,33–38] Data were requested from the authors of studies who did not report the accuracy data for the outcomes CIN2+ and CIN3+, separately for the two triage groups.[26,33–38] All of them provided the requested data and were offered coauthorship to this article. Data from the CLEAR trial, including triage of women aged 21 years or older with ASC-US, were downloaded from the file submitted by the manufacturer for FDA approval.[39]

Study characteristics

Three studies included women with ASC-US or LSIL identified through primary screening and [34,37,39] four studies recruited subjects in colposcopy clinics where women were referred to because of prior cervical abnormalities,[26,33,35,36] whereas one study used archived liquid samples reported as LSIL from a cytology biobank with the outcomes assessed through the follow-up pathology registry[38] (see Table 1). Adjudicated cytology results were used to define the ASC-US and LSIL triage groups in the French FASE trial.[34] From the CLEAR trial, we considered only the outcomes from colposcopy-targeted biopsies as this type of verification was generally used in the other included studies.[39] In all studies, collected cervical cells were transferred in PreservCyt, a transport medium developed for ThinPrep liquid-based cytology (Hologic Incorporated, Bedford, MA). By extracting only data from women with ASC-US or LSIL, the spectrum of disease should be considered representative for the study question, with some reservation for a German study where the cytology reporting system did not allow perfect translation into Bethesda terminology.[33] Complete verification with an acceptably valid reference standard was systematically provided, and incorporation bias was avoided in all studies. The delay between triage testing and verification was not sufficiently documented in four studies,[33–35,37] and independence between HPV testing and verification was unclear in two studies.[26,34] As HC2 and APTIMA produce machine-based results, we can assume independent assessment even when not explicitly reported. Clinical data were not always provided,[26,38] although this criterion can be considered as captured at least partially by the clinical setting. Equivocal results and withdrawal of cases were not optimally reported in most of the studies (see Table 1).

Accuracy in triage of ASC-US

The sensitivity of APTIMA to triage ASC-US varied between 75 and 100% for the outcome CIN2+ and between 93% and 100% for outcome CIN3+. The specificity ranged from 20 to 81% and from 38% to 81% for CIN2+ and CIN3+, respectively (see Table 2). The pooled sensitivity and specificity of APTIMA was 95.7% (95% CI = 91.5–97.2%) and 56.4% (95% CI = 44.7–67.5%), respectively, for CIN2+ and 96.2% (95% CI = 91.7–98.3%) and 54.9% (95% CI = 43.5–65.9%), respectively, for CIN3+ (see Table 2 and Fig. 1).

APTIMA and HC2 showed similar sensitivity for CIN2+ and CIN3+ lesions (pooled ratios not significantly different from unity). However, the specificity of APTIMA was significantly higher (ratio: 1.19, 95% CI = 1.08–1.31 for outcome CIN2+ and ratio: 1.18, 95% CI = 1.08–1.29 for outcome CIN3+; see Fig. 2).

Accuracy in triage of LSIL

In triage of LSIL, the range of variation in sensitivity of the APTIMA test was 75–100% for outcome CIN2+ and 83–100% for outcome CIN3+, whereas for specificity, the ranges were 24–74% (CIN2+) and 23–73% (CIN3+). Considering CIN2+ as the outcome, the pooled sensitivity and specificity

Table 1. Study characteristics with evaluation of the quality of each study according to the QUADAS check list

Study	Country	Clinical setting	Number of subjects	QUADAS check list										
				1	2	3	4	5	6	7	8	9	10	11
Szarewski *et al.*, 2008[36]	UK (England)	Colposcopy clinic	104 borderline nuclear abnormalities	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	U
			615 mild dyskaryosis											
Dockter *et al.*, 2009[22,26]	USA	Colposcopy clinic	159 ASC-US	Y	Y	Y	Y	Y	Y	U	Y	N	N	Y
			143 LSIL											
Clad *et al.*, 2011[33]	Germany	Colposcopy clinic	13 ASC-US	U	Y	U	Y	Y	Y	Y	Y	Y	Y	U
			44 LSIL											
Cuschieri *et al.*, 2010[38]	UK (Scotland)	Registry-linked biobank	543 mild dyskaryosis	Y	Y	Y	Y	Y	Y	Y	Y	U	U	U
Wu *et al.*, 2010[37]	China	Primary screening	88 ASC-US	Y	Y	U	Y	Y	Y	Y	Y	Y	N	Y
			12 LSIL											
CLEAR, 2011[39]	USA	Primary screening	541 ASC-US	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U
Monsonego *et al.*, 2011[34]	France	Primary screening	119 ASC-US	Y	Y	U	Y	Y	Y	U	Y	Y	Y	Y
			206 LSIL											
Ratnam *et al.*, 2011[35]	Canada	Colposcopy clinic	419 ASC-US	Y	Y	U	Y	Y	Y	Y	Y	Y	U	U
			324 LSIL											

QUADAS items: (1) representative spectrum; (2) acceptable reference test; (3) acceptable delay between tests; (4) partial verification avoided; (5) differential verification avoided; (6) incorporation avoided; (7) standard reference blinded to tests; (8) index and comparator tests blinded; (9) relevant clinical information; (10) uninterpretable/equivocal test results reported and (11) withdrawals explained. Each quality item is judged with: Y (fulfilled, green), U (unclear, yellow) and N (not fulfilled, red).

Table 2. Range (minimal–maximal values) and pooled sensitivity and specificity (in %) of APTIMA and HC2 to detect CIN2+ or CIN3+ in women with ASC-US or LSIL

Triage group	Outcome	Parameter	APTIMA		HC2	
			Min–max	Pooled value	Min–max	Pooled value
ASC-US	CIN2+	Sensitivity	75–100	95.7 (91.5–97.2)	75–100	93.8 (89.8–96.2)
		Specificity	20–81	56.4 (44.7–67.5)	40–74	46.8 (40.6–53.1)
	CIN3+	Sensitivity	93–100	96.2 (91.7–98.3)	83–100	95.5 (90.9–97.8)
		Specificity	38–81	54.9 (43.5–65.9)	23–73	44.9 (33.3–57.1)
LSIL	CIN2+	Sensitivity	75–100	91.0 (85.2–94.7)	92–100	95.5 (91.1–97.8)
		Specificity	24–74	42.5 (33.3–52.3)	9–40	28.6 (21.7–36.6)
	CIN3+	Sensitivity	83–100	96.7 (91.4–98.9)	95–100	98.8 (CI: 96.2–100.0)
		Specificity	23–73	38.7 (30.5–47.6)	8–36	27.8 (18.4–39.8)

were 91.0% (95% CI = 85.2–94.7%) and 42.5% (95% CI = 33.3–52.3%), respectively, whereas for CIN3+, the estimates were 96.7% (95% CI = 91.4–98.9%) and 38.7% (95% CI = 30.5–47.6%), respectively (Table 2).

APTIMA was, on average, as sensitive as HC2 but substantially and significantly more specific for CIN2+ (ratio: 1.37; 95% CI = 1.22–1.54) as well as for CIN3+ (ratio: 1.35; 95% CI = 1.11–1.66) (Table 3 and Fig. 2).

Influence of study characteristics

None of the study characteristics or QUADAS quality items influenced significantly the relative accuracy parameters, with the exception of one situation. In LSIL triage, the relative specificity was not significantly different from unity in the Scottish study where women had repeated LSIL and where samples were selected from a cytological archive.[38]

Variation of the accuracy by age group was documented in only one study. In the CLEAR trial, specificity of ASC-US triage increased by age for both tests. However, the relative specificity of APTIMA compared to HC2 to exclude CIN2+ decreased by age: 1.27 (1.05–1.53) in age group 21–29 years, 1.14 (0.97–1.27) in age group 30–39 years and 1.04 (0.95–1.13) in women aged 40 years or older.

Discussion

Our meta-analysis showed greater specificity without loss in sensitivity for the new APTIMA HPV assay in triage of ASC-

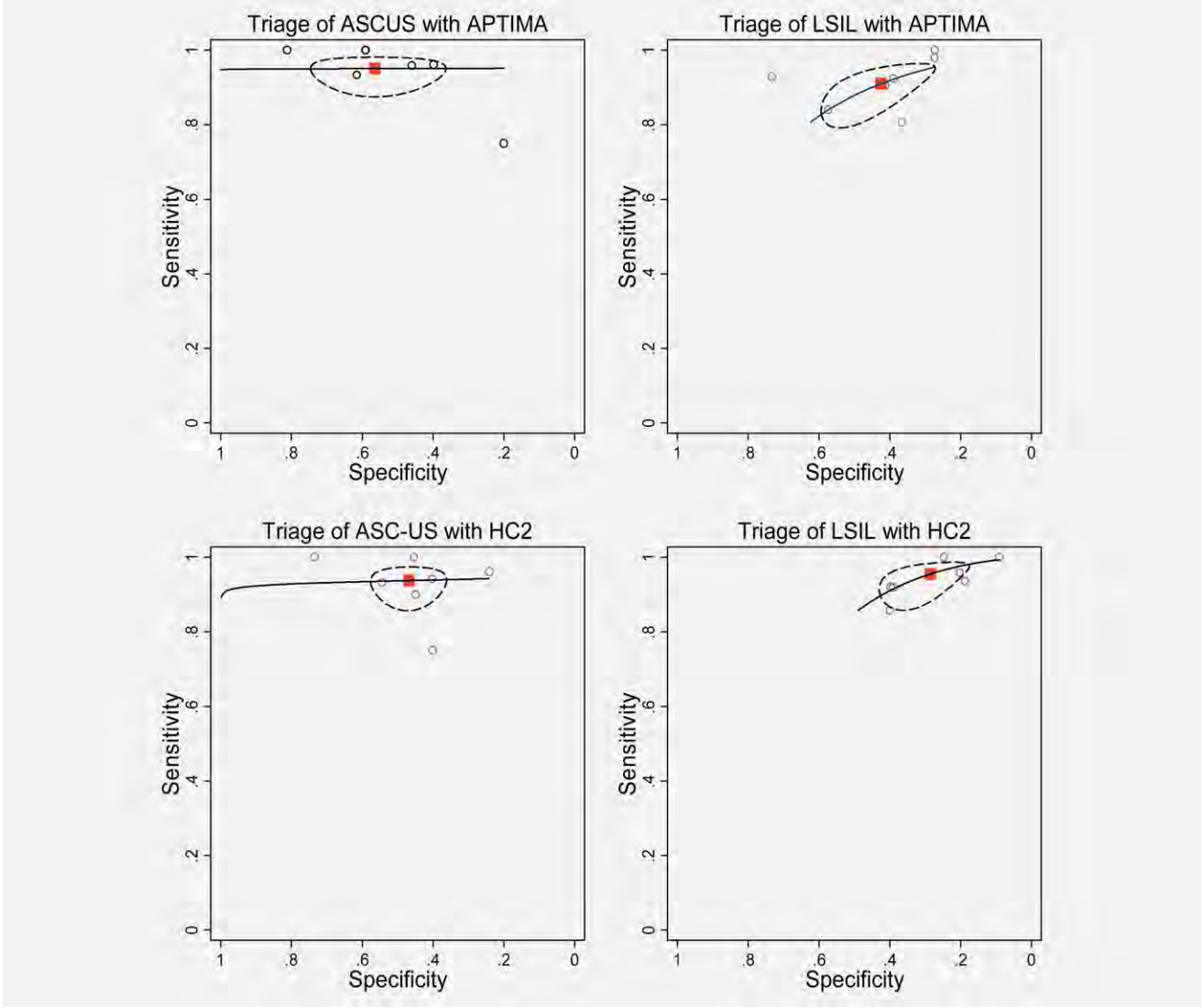


Figure 1. Meta-analysis of the sensitivity and specificity of APTIMA (top) and HC2 (bottom) in triage of women with ASC-US (left) or LSIL (right) to detect CIN2+. Square: summary point; small circles: individual studies; continuous line: SROC curve; interrupted line: 95% confidence ellipse.

US and LSIL with respect to the detection of high-grade CIN compared to HC2. HC2 is currently considered as a validated test recommended internationally discriminating between women with ASC-US needing further diagnostic work-up and those who can be called back for routine screening.[2,44] However, findings on the performance of HPV testing in triage of LSIL, in particular regarding its specificity, are divergent, which are mainly due to local differences in cytological interpretation as reflected by the variation of HPV-positivity rate.[45] Reflex HPV testing with HC2 in case of LSIL is therefore not generally recommended.[1,5] An assay with as good sensitivity but with higher specificity than HC2 is required that could be proposed in triage of LSIL. The APTIMA test seems to be such a test with a specificity for excluding high-grade CIN that is 37% (95% CI = 22–54%) higher than HC2 and a sensitivity that is not lower.

Only eight studies have been conducted up to now comparing HC2 and APTIMA and providing separate data for triage of ASC-US or LSIL. However, one study evaluated the accuracy of APTIMA in LSIL, but did not include a comparator test.[46] Adding this report, comprising 405 women with LSIL, to the meta-analysis changed the pooled accuracy parameters by less than 0.3%. Two other studies assessed the diagnostic performance in women with abnormal cytology without separation by cytological category and using other comparator assays. Ovestad *et al.*[47] compared APTIMA with two other high-risk HPV DNA tests (Amplicor, Roche Diagnostics GmbH, Mannheim, Germany and COBAS-4800, Roche Molecular Diagnostics, Pleasanton, CA) and with one mRNA test (PreTect HPV Proofer, NorChip AS, Klokkarstua, Oslo, Norway) in a series of women with ASC-US or LSIL without separation. The sensitivity of APTIMA for CIN2+

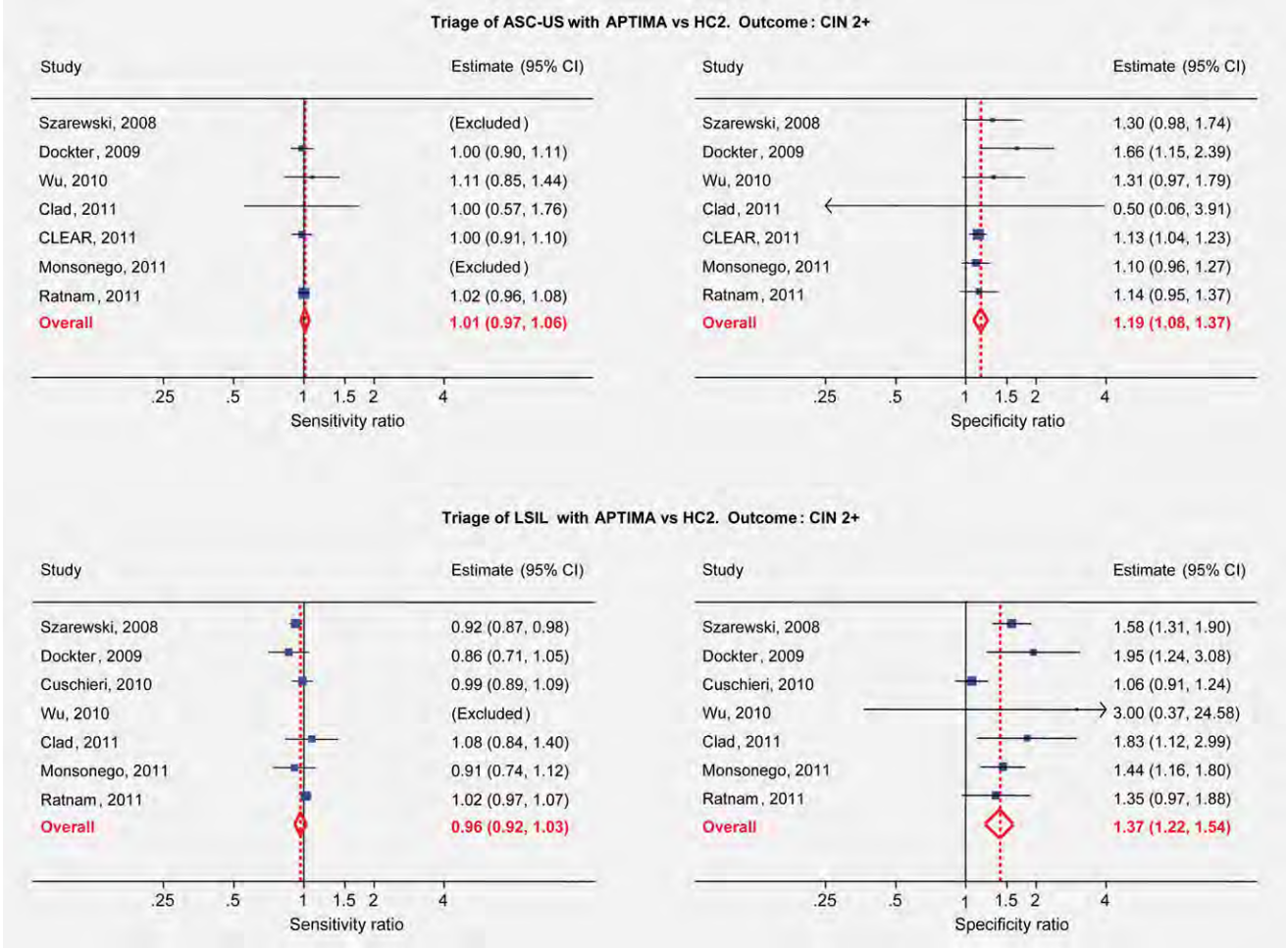


Figure 2. Meta-analysis of the relative sensitivity (left) and specificity (right) of APTIMA *versus* HC2 to exclude the presence of CIN2+ in triage of ASC-US (top) or LSIL (bottom). Studies where both tests showed 100% sensitivity were excluded from the meta-analysis.

Table 3. Pooled relative sensitivity and specificity of APTIMA compared to HC2

Triage group	Outcome	Parameter	Ratio (APTIMA/HC2)
ASC-US	CIN2+	Sensitivity	1.01 (0.97–1.06)
		Specificity	1.19 (1.08–1.31)
	CIN3+	Sensitivity	1.01 (0.96–1.06)
		Specificity	1.18 (1.08–1.29)
LSIL	CIN2+	Sensitivity	0.96 (0.92–1.03)
		Specificity	1.37 (1.22–1.54)
	CIN3+	Sensitivity[1]	0.98 (0.91–1.06)
		Specificity[1]	1.35 (1.11–1.66)

[1]The *metadas* SAS macro failed to converge. Therefore, the pooled relative sensitivity and specificity were computed separately as ratios of two proportions using a random effect model.

was similar to that of the other two HPV DNA tests; however, APTIMA was substantially more specific than Amplicor (ratio: 2.14; 95% CI = 1.23–2.73). Compared to PreTect, a test that detects mRNA of five HPV types, APTIMA was substantially more sensitive (ratio: 1.91; 95% CI = 1.43–2.56) but less specific (ratio: 0.47; 95% CI = 0.34–0.63). Castle *et al.*[41] compared a prototype version of APTIMA at two different cutoffs for test positivity with detection of 14 high-risk HPV genotypes with the Linear Array test (Roche Diagnostics GmbH, Mannheim, Germany) in women with diverse abnormal cytology. Although these studies did not contribute to our meta-analysis, they seem to corroborate the statement that APTIMA is as sensitive as other currently used HPV DNA test systems for the detection of CIN2+ but often more specific. Moreover, the FASE and CLEAR trials, contributing data to our meta-analysis, also demonstrated similar sensitivity but higher specificity of APTIMA compared to HC2 in a primary screening setting.[34,39]

Because APTIMA detects mRNA, which is more fragile than DNA, its accuracy could be compromised by conditions of sample transport and storage. However, APTIMA's sensitivity did not seem to be affected in the Scottish study where stored cellular material from a biobank was used[38]; however, in contrast to the other studies, the specificity here was not higher than HC2. Probably the lack of improved specificity in the Scottish study can be ascribed to the fact that all LSIL

cases showed persistent cytological abnormality. All studies used the PreservCyt liquid as transport medium. More data are needed to assess the possible influence of other transport media for liquid-based cytology.

The higher specificity of APTIMA may be ascribed to reduced cross-reaction with nontargeted low-risk or intermediate risk types (such as HPV53), which has been reported for HC2.[48] However, we cannot exclude that the increased specificity was partly due to small missed CIN2+ cases, not picked up by colposcopy and colposcopy-targeted punch biopsy, which were HC2 positive but APTIMA negative.

Clinical implications

Given widely accepted evidence for the use of HC2 in triage of ASC-US and the similar sensitivity and higher specificity of APTIMA to detect high-grade CIN, we certainly can recommend APTIMA for ASC-US triage.

As HC2 is not generally recommended in triage of LSIL, a more formal judgment of the quality of evidence derivable from our systematic review is warranted based on the following five criteria: (*i*) consistency, (*ii*) large and precise effect size, (*c*) directness, (*d*) study design and (*e*) absence of publication bias.[49,50] Although wide variability in absolute accuracy, in particular, in specificity, of both tests, the relative accuracy estimates were consistent (sensitivity ratio never significantly different from unity, specificity ratio always >1 and significantly >1 in four studies; see Fig. 2).[38] Moreover, the magnitude of the difference in specificity was substantial (pooled absolute specificity of 43% for APTIMA *versus* 29% for HC2, pooled ratio 1.37, with the confidence interval clearly excluding 1). Verification with a reference standard was performed in all cases, and the QUADAS evaluation was considered satisfactory. Moreover, the accuracy could be evaluated separately for ASC-US and LSIL cases making the evidence direct. Statistical tests for publication bias were insignificant (data not shown). Despite the limited number of studies and the low power of the statistical tests for publication bias, we could restrict the risk of publication bias by consulting other sources such as conference reports and the websites of regulatory agencies. Given these arguments, we conclude that APTIMA can also be recommended to triage women with LSIL.

Conclusions

The APTIMA HPV test is as sensitive as HC2 to detect CIN2+ or CIN3+ in women with ASC-US or LSIL. The specificity of APTIMA is on average 19% (95% CI = 8–29%) more specific in triage of ASC-US and 37% (95% CI = 22–54%) more specific in triage of LSIL. The APTIMA HPV test might be considered for use not only for triage of women with ASC-US but also for women with LSIL to identify those who need further diagnostic assessment with colposcopy and biopsy.

References

1. Jordan J, Arbyn M, Martin-Hirsch P, et al. European guidelines for quality assurance in cervical cancer screening: recommendations for clinical management of abnormal cervical cytology, Part 1. *Cytopathology* 2008;19:342–54.
2. Partridge EE, Abu-Rustum NR, Campos SM, et al. Cervical cancer screening: NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2010;8:1358–86.
3. Stampler KM, Dunton CJ. Cervical neoplasia guidelines: United States and Europe compared. *J Low Genit Tract Dis* 2010;14:142–7.
4. Wright TC, Jr, Massad LS, Dunton CJ, et al. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. *Am J Obstet Gynecol* 2007;197:346–55.
5. Solomon D, Papillo JL, Davey DD. Statement on HPV DNA test utilization. *Am J Clin Pathol* 2009;131:768–9.
6. Sorbye SW, Fismen S, Gutteberg T, et al. Triage of women with minor cervical lesions: data suggesting a "test and treat" approach for HPV E6/E7 mRNA testing. *PLoS One* 2010;5:e12724 (1)–e12724 (6).
7. TOMBOLA Group. Cytological surveillance compared with immediate referral for colposcopy in management of women with low grade cervical abnormalities: multicentre randomised controlled trial. *BMJ* 2009;339:b2546–b2557.
8. Kelly RS, Patnick J, Kitchener HC, et al. HPV testing as a triage for borderline or mild dyskaryosis on cervical cytology: results from the Sentinel Sites study. *Br J Cancer* 2011;105:983–8.
9. Sorbye SW, Arbyn M, Gutteberg T, et al. Triage of women with low-grade cervical lesions—HPV mRNA testing versus repeat cytology. *PLoS One* 2011;6:e24083-1–e24083-6.
10. Melnikow J, Nuovo J, Willan AR, et al. Natural history of cervical squamous intraepithelial lesions: a meta-analysis. *Obstet Gynecol* 1998;92: 727–35.
11. Holowaty P, Miller AB, Rohan T, et al. Natural history of dysplasia of the uterine cervix. *J Natl Cancer Inst* 1999;91:252–8.
12. Lorincz AT. Hybrid capture. *Clin Chem* 1998;44: 1363.
13. Arbyn M, Buntinx F, Van Ranst M, et al. Virologic versus cytologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. *J Natl Cancer Inst* 2004;96:280–93.
14. Arbyn M, Sasieni P, Meijer CJ, et al. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. *Vaccine* 2006;24 (Suppl. 3): S78–S89.
15. Cuzick J, Arbyn M, Ronco G, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. *Vaccine* 2008;26 (Suppl. 10):K29–K41.
16. ASCUS-LSIL Triage Study Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. *Am J Obstet Gynecol* 2003;188:1383–92.
17. ASCUS-LSIL Triage Study Group. A randomized trial on the management of low-grade squamous intraepithelial lesion cytology interpretations. *Am J Obstet Gynecol* 2003;188:1393–400.
18. Arbyn M, Paraskevaidis E, Martin-Hirsch P, et al. Clinical utility of HPV DNA detection: triage of minor cervical lesions, follow-up of women treated for high-grade CIN. An update of pooled evidence. *Gynecol Oncol* 2005;99 (Suppl. 3):7–11.
19. Arbyn M, Martin-Hirsch P, Buntinx F, et al. Triage of women with equivocal or low-grade cervical cytology results. A meta-analysis of the HPV test positivity rate. *J Cell Mol Med* 2009;13: 648–59.
20. Ronco G, Cuzick J, Segnan N, et al. HPV triage for low grade (L-SIL) cytology is appropriate for women over 35 in mass cervical cancer screening using liquid based cytology. *Eur J Cancer* 2007; 43:476–80.
21. Castle PE, Fetterman B, Thomas CJ, et al. The age-specific relationships of abnormal cytology and human papillomavirus DNA results to the risk of cervical precancer and cancer. *Obstet Gynecol* 2010;116:76–84.
22. Dockter J, Schroder A, Eaton B, et al. Analytical characterization of the APTIMA HPV assay. *J Clin Virol* 2009;45 (Suppl. 1):S39–S47.
23. Lundberg GD, National Cancer Institute. The 1988 Bethesda System for reporting cervical/ vaginal cytologic diagnoses. *JAMA* 1989;262: 931–4.
24. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002;287: 2114–19.
25. Evans DMD, Hudson EA, Brown CL, et al. Terminology in gynaecological cytopathology: report of the working party of the BSCC. *J Clin Pathol* 1986;39:933–44.

26. Dockter J, Schroder A, Hill C, et al. Clinical performance of the APTIMA HPV assay for the detection of high-risk HPV and high-grade cervical lesions. *J Clin Virol* 2009;45:S55–S61.
27. Whiting P, Rutjes AWS, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:1–13.
28. Dersimonian R, Laird NM. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986;7: 177–88.
29. Harbord RM, Deeks JJ, Egger M, et al. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;8:239–51.
30. Harbord RM, Whiting P. Metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J* 2009;9:211–29.
31. Takwoingi Y, Diagnostic Test Accuracy Working Group. METADAS: a SAS macro for meta-analysis of diagnostic accuracy studies. Available from: http://srdta.cochrane.org/sites/srdta.cochrane.org/files/uploads/METADAS%20-Quick%20Reference%20WorkedExample%20v1.3.pdf. Accessed October 31, 2011.
32. Diagnostic Test Accuracy Working Group. Handbook for diagnostic test accuracy reviews. Available from: http://srdta.cochrane.org/handbook-dta-reviews. 2011, Accessed October 31, 2011.
33. Clad A, Reuschenbach M, Weinschenk J, et al. Performance of the APTIMA high-risk HPV mRNA assay in a referral population in comparison with Hybrid Capture 2 and cytology. *J Clin Microbiol* 2011;49:1071–6.
34. Monsonego J, Hudgens MG, Zerat L, et al. Evaluation of oncogenic human papillomavirus RNA and DNA tests with liquid based cytology in primary cervical cancer screening (The FASE study). *Int J Cancer* 2011;129:691–701.
35. Ratnam S, Coutlee F, Fontaine D, et al. Aptima HPV E6/E7 mRNA test is as sensitive as Hybrid Capture 2 Assay but more specific at detecting cervical precancer and cancer. *J Clin Microbiol* 2011;49:557–64.
36. Szarewski A, Ambroisine L, Cadman L, et al. Comparison of predictors for high-grade cervical intraepithelial neoplasia in women with abnormal smears. *Cancer Epidemiol Biomarkers Prev* 2008; 17:3033–43.
37. Wu R, Belinson SE, Du H, et al. Human papillomavirus messenger RNA assay for cervical cancer screening: the Shenzhen Cervical Cancer Screening Trial I. *Int J Gynecol Cancer* 2010;20: 1411–14.
38. Cuschieri K, Kitchener H, Horne AW, et al. Clinical performance of a broad-spectrum RNA assay in a high-prevalence disease setting. In 26th international Papillomavirus Conference and Clinical and Public Health Workshops, Montreal, 2010.
39. Gen-Probe. APTIMA HPV Assay. Available from: http://www.accessdata.fda.gov/cdrh_docs/pdf10/P100042c.pdf. 2011:-502170 rev. A. 502170 rev. A. Accessed November 30, 2011.
40. Getman D, Aiyer A, Dockter J, et al. Efficiency of the APTIMA HPV assay for detection of HPV RNA and DNA targets. *J Clin Virol* 2009;45 (Suppl. 1):S49–S54.
41. Castle PE, Dockter J, Giachetti C, et al. A cross-sectional study of a prototype carcinogenic human papillomavirus E6/E7 messenger RNA assay for detection of cervical precancer and cancer. *Clin Cancer Res* 2007;13:2599–605.
42. Cuzick J, Ambroisine L, Cadman L, et al. Performance of the Abbott RealTime high-risk HPV test in women with abnormal cervical cytology smears. *J Med Virol* 2010;82:1186–91.
43. Reuschenbach M, Clad A, Knebel Doeberitz C, et al. Performance of p16INK4a-cytology, HPV mRNA, and HPV DNA testing to identify high grade cervical dysplasia in women with abnormal screening results. *Gynecol Oncol* 2010; 119:98–105.
44. European Commission. European guidelines for quality assurance in cervical cancer screening. Luxembourg: Office for Official Publications of the European Communities, 2008. 291.
45. Arbyn M, Roelens J, Martin-Hirsch P, et al. Use of HC2 to triage women with borderline and mild dyskaryosis in the UK. *Br J Cancer* 2011; 105:877–80.
46. Waldstrom M, Ornskov D. Clinical performance of a human papillomavirus messenger RNA test (APTIMA HPV assay) on residual material from archived 3-year-old preservcyt samples with low-grade squamous intraepithelial lesion. *Arch Pathol Lab Med* 2011;135:1052–6.
47. Ovestad IT, Vennestrom U, Andersen L, et al. Comparison of different commercial methods for HPV detection in follow-up cytology after ASCUS/LSIL, prediction of CIN2–3 in follow up biopsies and spontaneous regression of CIN2–3. *Gynecol Oncol* 2011;123:278–83.
48. Castle PE, Solomon D, Wheeler CM, et al. Human papillomavirus genotype specificity of Hybrid Capture 2. *J Clin Microbiol* 2008;46: 2595–604.
49. Atkins D, Best D, Briss PA, et al, GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328: 1490.
50. Schunemann HJ, Oxman AD, Brozek J, et al. GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;336:1106–10.

Appendix: Search Strategy for Retrieval of Studies Allowing the Evaluation of APTIMA in Triage of ASC-US or LSIL

#1: Cervix OR cervic*
#2: Cancer OR carcinoma OR neoplas* OR dysplas* OR CIN[tw] OR CINII*[tw] OR CIN2*[tw] OR CINIII*[tw] OR CIN3[tw] OR SIL[tw] OR LSIL[tw] OR L-SIL OR ASC-US[tw] OR "ASC-US"[tw] OR "low grade" OR low-grade OR mild OR equivocal OR borderline.
#3: #1 AND #2.
#4: diagnos* OR test OR assay
#5: HPV OR "human papillomavirus" OR papillomavir*
#6: APTIMA OR RNA OR "ribonucleic acid" OR transcript*
#7: #4 AND #5 AND #6
#8: #3 AND #7
#9: screening OR triage OR management OR follow-up
#10: #8 AND #9
#11: Limits: Humans, Female
#12: Publication Date from January 1, 2007 to October 29, 2011
#13: #10 AND #11 AND #12

Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis

Marc Arbyn, Freija Verdoodt, Peter J F Snijders, Viola M J Verhoef, Eero Suonio, Lena Dillner, Silvia Minozzi, Cristina Bellisario, Rita Banzi, Fang-Hui Zhao, Peter Hillemanns, Ahti Anttila

Lancet Oncol 2014; 15: 172–83

Published Online
January 14, 2014
http://dx.doi.org/10.1016/S1470-2045(13)70570-9

See Comment page 128

Unit of Cancer Epidemiology, Scientific Institute of Public Health, Brussels, Belgium (M Arbyn DrTMH, F Verdoodt PhD); **Department of Pathology, VU University Medical Center, Amsterdam, Netherlands** (Prof P J F Snijders PhD, V M J Verhoef MD); **International Agency for Research on Cancer, Lyon, France** (E Suonio MD); **Karolinska Institute, Stockholm, Sweden** (L Dillner PhD); **Unit of Cancer Epidemiology, Department of Oncology, Piedmont Centre for Cancer Prevention, S Giovanni University Hospital, Turin, Italy** (S Minozzi MD, C Bellisario MSc); **IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy** (R Banzi PhD); **Department of Cancer Epidemiology, Cancer Institute and Hospital, Chinese Academy of Medical Sciences, Beijing, China** (Prof F-H Zhao PhD); **Department of Gynaecology and Obstetrics, Hannover Medical School, Hannover, Germany** (Prof P Hillemanns PhD); **and Finnish Cancer Registry, Helsinki, Finland** (A Anttila PhD)

Correspondence to:
Dr Marc Arbyn, Unit of Cancer Epidemiology, Scientific Institute of Public Health, J Wytsmanstreet 14, B1050 Brussels, Belgium
marc.arbyn@wiv-isp.be

Summary

Background Screening for human papillomavirus (HPV) infection is more effective in reducing the incidence of cervical cancer than screening using Pap smears. Moreover, HPV testing can be done on a vaginal sample self-taken by a woman, which offers an opportunity to improve screening coverage. However, the clinical accuracy of HPV testing on self-samples is not well-known. We assessed whether HPV testing on self-collected samples is equivalent to HPV testing on samples collected by clinicians.

Methods We identified relevant studies through a search of PubMed, Embase, and CENTRAL. Studies were eligible for inclusion if they fulfilled all of the following selection criteria: a cervical cell sample was self-collected by a woman followed by a sample taken by a clinician; a high-risk HPV test was done on the self-sample (index test) and HPV-testing or cytological interpretation was done on the specimen collected by the clinician (comparator tests); and the presence or absence of cervical intraepithelial neoplasia grade 2 (CIN2) or worse was verified by colposcopy and biopsy in all enrolled women or in women with one or more positive tests. The absolute accuracy for finding CIN2 or worse, or CIN grade 3 (CIN3) or worse of the index and comparator tests as well as the relative accuracy of the index versus the comparator tests were pooled using bivariate normal models and random effect models.

Findings We included data from 36 studies, which altogether enrolled 154 556 women. The absolute accuracy varied by clinical setting. In the context of screening, HPV testing on self-samples detected, on average, 76% (95% CI 69–82) of CIN2 or worse and 84% (72–92) of CIN3 or worse. The pooled absolute specificity to exclude CIN2 or worse was 86% (83–89) and 87% (84–90) to exclude CIN3 or worse. The variation of the relative accuracy of HPV testing on self-samples compared with tests on clinician-taken samples was low across settings, enabling pooling of the relative accuracy over all studies. 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. By contrast, some PCR-based HPV tests generally showed similar sensitivity on both self-samples and clinician-based samples.

Interpretation 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.

Funding The 7th Framework Programme of the European Commission, the Belgian Foundation against Cancer, the International Agency for Research on Cancer, and the German Guideline Program in Oncology.

Introduction

In view of the strong aetiological link between persistent infection with high-risk types of the human papillomavirus (HPV) and the development of cervical cancer, assays have been developed to detect nucleic acid sequences of the virus.[1,2] Meta-analyses have provided clinical evidence that has led to widely accepted recommendations to use HPV tests to triage women with equivocal cervical cytology and to predict recurrence after treatment of cervical precancer.[3–6] Data from randomised trials have consistently shown that women with a prior negative HPV test have a lower risk of developing grade 3 cervical intraepithelial neoplasia (CIN) and invasive cervical cancer compared with women with a prior normal Pap smear.[7] This new evidence underpins the recommendation to change the policy of secondary prevention of cervical cancer and to use an HPV assay as the primary screening test, used either alone instead of a Pap smear or together with a Pap test.[8,9] Moreover, HPV testing can be done on a vaginal sample taken by the women themselves, which might offer opportunities to reach those who are reluctant to undergo gynaecological examinations.[10,11]

Previous systematic reviews have summarised the performance of HPV testing on self-samples but these reviews were done 6–8 years ago and included mainly small studies, they assessed only virological outcomes, or they did not compare the accuracy for high-grade CIN

with tests on clinician-taken samples.[12–14] A review drew attention to the need for a comprehensive updated meta-analysis in view of the large amount of new data from large studies that have used a wide range of tests and collection devices.[10]

In this meta-analysis, we assessed the clinical accuracy of HPV testing on self-samples to detect underlying high-grade CIN or cancer. Moreover, we compared the accuracy of HPV testing in self-samples with that of HPV testing and cytology processing on samples taken by a clinician. We did not assess cytological processing of self-samples because it was previously shown to be poorly sensitive for picking up high-grade CIN lesions.[10,15,16]

We focused on primary screening for cervical cancer, because self-sampling is most often used in this type of setting and because the absolute accuracy might be different in women who are followed up because of prior cervical abnormalities. However, the relative accuracy of HPV testing in a sample taken by a woman compared with a sample taken by a health professional could be similar across settings. Assessment of the relative accuracy allowed us to include more studies, incorporate randomised trials, and increase the power for explaining heterogeneity about, in particular, the effects of collection devices and HPV assays. We aimed to find out whether an HPV test on a self-sample is as good as a test on a sample taken by a clinician in women attending cervical cancer screening.

Methods

Search strategy and selection criteria

We searched PubMed, Embase, and CENTRAL for eligible studies (see appendix for the clinical questions and search terms used). We searched for papers published between Jan 1, 1990, and June 3, 2013. We also used Scopus to investigate citations of previous systematic reviews on HPV testing of self-samples,[10,12–14] and the reference lists of selected references. Additionally, we searched for unpublished reports in the abstract books of the three most recent international conferences of the Papillomavirus Society (Montreal, July 3–8, 2010; Berlin, Sept 17–22, 2011, and San Juan, Nov 30–Dec 6, 2012). We applied no language restrictions.

Studies were eligible for inclusion if the following criteria were fulfilled: a vaginal sample was self-taken by a woman (self-sample) followed by 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 on both samples or the clinician-taken sample was examined microscopically for presence of cytological epithelial lesions; and 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. Studies with cytological follow-up for women with negative colposcopy test results at baseline assessment were accepted as fulfilling the third selection criterion but were indexed for sensitivity analyses.

We contacted authors to provide accuracy data separately for the outcome CIN2 or worse when only the outcome CIN grade 3 (CIN3) or worse was reported.

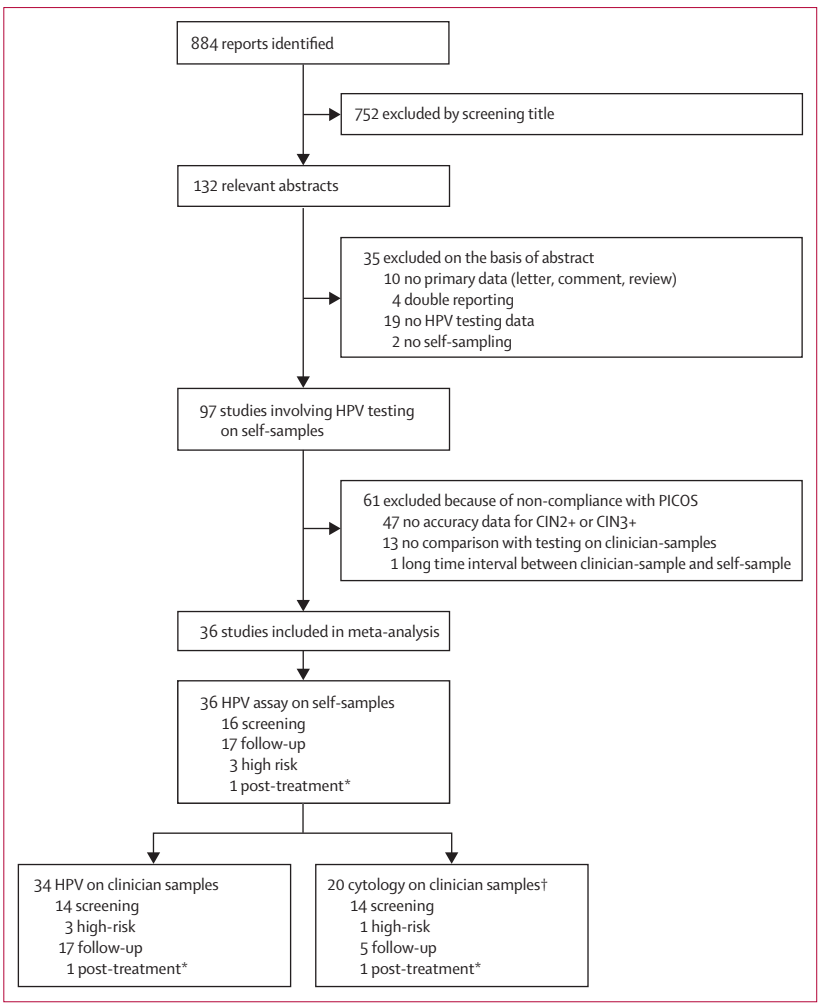


Figure 1: **Study flow chart**
CIN2 or worse=cervical intraepithelial neoplasia of grade 2 or worse. CIN3 or worse=cervical intraepithelial neoplasia of grade 3 or worse. HPV=human papillomavirus. PICOS=Population/Index/Comparator/Outcome/Study. *One study[48] included women in a follow-up setting and a post-treatment surveillance setting. †18 of 20 studies included both cytology and HPV testing on the clinician-taken samples, whereas two trials[47,50] included only cytology on the clinician-taken samples.

See **Online** for appendix

Clinical questions and data extraction

We aimed to answer two questions. The first was to establish the absolute accuracy of HPV testing on a self-sample (index test) and the accuracy of cytological processing or HPV testing on a cervical cell specimen taken by a clinician (comparator tests). The second was to establish the relative accuracy of HPV testing on a self-sample compared with the comparator tests on a clinician-taken sample. Accuracy was determined for the

	Patient selection		Index and comparator tests		Reference test			Flow and timing						Concerns of applicability: risk of bias		
	P1	P2	T1	T2	R1	R2	R3	F1	F2	F3	F4	F5	F6	Patient selection	Index and comparator test	Reference test
Morrison et al, 1992[30]	U	U	U	U	Y	U	Y	Y	Y	Y	Y	Y	Y	Moderate	Moderate	Low
Hillemanns et al, 1999[31]	U	U	U	U	Y	U	Y	U	Y	Y	U	N	N	Moderate	Moderate	Low
Sellors et al, 2000[32]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Wright et al, 2000[33]	U	U	Y	Y	Y	Y	Y	N	N	Y	N	N	N	Moderate	Low	Low
Belinson et al, 2001[34]	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Lorenzato et al, 2002[35]	Y	Y	Y	U	Y	U	Y	Y	Y	Y	Y	Y	N	Low	Low	Low
Nobbenhuis et al, 2002[16]	Y	U	U	U	Y	U	Y	Y	Y	Y	Y	Y	Y	Low	Moderate	Low
Garcia et al, 2003[15]	Y	U	Y	U	Y	Y	Y	Y	Y	Y	N	Y	N	Low	Low	Low
Salmeron et al, 2003[36]	U	U	Y	U	Y	N	Y	U	N	Y	Y	Y	Y	Moderate	Low	Moderate
Brink, 2006[37]	U	U	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	Moderate	Low	Low
Daponte et al, 2006[38]	Y	U	U	U	Y	U	Y	U	Y	Y	U	N	N	Low	Moderate	Low
Girianelli et al, 2006[39]	U	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Moderate	Low	Low
Holanda et al, 2006[40]	U	U	Y	U	Y	Y	Y	Y	Y	Y	Y	N	N	Moderate	Low	Low
Seo et al, 2006[41]	U	U	Y	U	Y	U	Y	Y	Y	Y	N	N	N	Moderate	Low	Low
Szarewski et al, 2007[42]	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Qiao et al, 2008[43]	Y*	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Bhatla et al, 2009[44]	Y	U	Y	U	Y	U	Y	Y	Y	Y	Y	N	Y	Low	Low	Low
Balasubra et al, 2010[45]	Y	U	Y	U	Y	U	Y	Y	Y	Y	Y	Y	Y	Low	Low	Low
Gustavsson et al, 2011[46]	U	U	U	U	Y	Y	Y	Y	Y	Y	Y	N	N	Moderate	Moderate	Low
Lazcano-P, 2011[47]	Y	U	Y	U	Y	U	Y	U	N	Y	Y	N	N	Low	Low	Low
Taylor et al, 2011[48]	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	Low	Low	Low
Twu et al, 2011[49]	U	U	Y	U	Y	U	Y	U	N	Y	N	Y	Y	Moderate	Low	Low
Wikstrom et al, 2011[50]	Y	U	Y	U	Y	U	Y	N	N	Y	Y	N	N	Low	Low	Low
Belinson et al, 2012[51]	U	U	Y	U	Y	U	Y	Y	N	Y	Y	N	Y	Moderate	Low	Low
Dijkstra et al, 2012[52]	U	U	U	Y	Y	U	Y	Y	Y	Y	N	N	N	Moderate	Low	Low
Longatto-F et al, 2012[53]	Y	U	Y	Y	Y	N	Y	U	Y	Y	Y	U	U	Low	Low	Moderate
van Baars et al, 2012[54]	U	U	U	U	N†	U	U	Y	N	U	Y	N	N	Moderate	Moderate	High
Zhao et al, 2012‡[55]	Y	Y	Y	U	Y	U	Y	Y	Y	Y	Y	N	N	Low	Low	Low
Darlin et al, 2013[56]	U	U	U	U	Y	U	Y	Y	Y	Y	Y	Y	N	Moderate	Moderate	Low
Geraets et al, 2013[57]	Y	U	U	U	Y	U	Y	Y	Y	Y	N	N	N	Low	Moderate	Low
Guan et al, 2013[58]	N	U	U	U	Y	U	Y	Y	Y	Y	N	N	N	High	Moderate	Low
Jentschke et al, 2013a[60]	U	U	U	U	Y	U	Y	U	Y	Y	Y	N	N	Moderate	Moderate	Low
Jentschke et al, 2013b[59]	U	U	Y	U	Y	U	Y	U	Y	Y	Y	N	N	Moderate	Low	Low
Nieves et al, 2013[61]	Y	Y	U	Y	Y	N	Y	U	Y	Y	Y	Y	Y	Low	Low	Moderate

QUADAS items:[19] P1=acceptable enrolment method, P2=inappropriate exclusions avoided, T1=prespecified test cut-off, T2=results of the index test and the comparator are masked towards each other and both the index and comparator tests are masked towards the reference test, R1=acceptable reference test, R2=results of the reference test are masked towards the index and comparator tests, R3=incorporation bias avoided, F1=acceptable delay between triage tests and reference test, F2=partial verification avoided, F3=differential verification avoided, F4=withdrawals explained, F5=uninterpretable results reported for tests, F6=uninterpretable results reported for reference test. Each quality item is judged with the following: Y=fulfilled, U=unclear, and N=not fulfilled. *Participants recruited from randomly selected communes. †Only 44/134 women were verified histologically, the rest of the participants were verified cytologically. ‡Only data for SPOCCS III-1, III-2, and III-3 are included in this meta-analysis; data also reported jointly by Belinson et al.[64]

Table 1: **Quality assessment of all included studies**

disease outcomes CIN2 or worse or CIN3 or worse. We analysed the following groups: those attending routine cervical cancer screening, high-risk women, and those referred to colposcopy because of previous positive screening results. We designed a protocol in accordance with PRISMA guidelines for reporting of meta-analyses.[17,18]

Eligibility of studies for inclusion was checked independently by at least two investigators and, subsequently, the numbers of the true and false positives and negatives for each combination of sampling, test, and disease outcome were extracted for all studies which had complete verification of test results with a reference standard. For randomised trials comparing self-samples with clinician-taken samples in which verification was restricted to patients being positive for one test, we extracted only the number of true-positives in those screened. Data from these trials was only used in the pooling of the relative sensitivity. Information about study participants, setting, tests, sampling

	Test cutoff	Number of studies (number of test-device combinations)		Sensitivity (%; 95% CI)		Specificity (%; 95% CI)	
		CIN2 or worse	CIN3 or worse	CIN2 or worse	CIN3 or worse	CIN2 or worse	CIN3 or worse
Primary screening							
Self-sample							
HPV	As defined by manufacturer	14 (16)	6 (8)	76% (69–82)	84% (72–92)	86% (83–89)	87% (84–90)
Clinician-taken sample							
HPV	As defined by manufacturer	14 (16)	6 (8)	91% (87–94)	95% (91–97)	88% (85–91)	89% (87–92)
Cytology	ASC-US or worse	12	6	83% (75–89)	91% (85–95)	91% (87–94)	89% (86–91)
	LSIL or worse	8	5	71% (66–76)	78% (72–85)*	97% (97–98)	97% (96–97)*
Screening of high-risk group							
Self-sample							
HPV	As defined by manufacturer	3 (4)	1	75% (58–87)	42% (27–57)†	86% (77–92)	81% (76–87)†
Clinician-taken sample							
HPV	As defined by manufacturer	3 (4)	1	88% (78–93)	80% (67–93)†	88% (81–93)	82% (77–88)†
Cytology	ASC-US or worse	1	0	77% (64–91)†	..	87% (84–90)†	..
	LSIL or worse	1	0	70% (55–85)†	..	95% (93–97)†	..
Follow-up							
Self-sample							
HPV	As defined by manufacturer	17 (19)	5 (7)	84% (78–89)	85% (76–91)	56% (49–63)	45% (36–54)
Clinician-taken sample							
HPV	As defined by manufacturer	17 (19)	5 (7)	91% (86–94)	96% (92–1 00)*	58% (48–67)	46% (35–57)*
Cytology	ASC-US or worse	5	0	85% (77–91)	..	69% (57–80)	..

ASC-US=atypical squamous cells of undetermined significance. CIN2=cervical intraepithelial neoplasia of grade 2. CIN3=cervical intraepithelial neoplasia of grade 3. HPV=assay that picks up high-risk types of the human papillomavirus. LSIL=low-grade squamous intraepithelial lesion. *Sensitivity and specificity pooled separately because lack of fit of the bivariate normal model. †No meta-analytical pooling because only one study.

Table 2: **Absolute sensitivity and specificity of human papillomavirus testing on self-samples, and human papillomavirus testing and cytology on clinician-taken samples, by clinical setting and grade of cervical intraepithelial neoplasia**

devices, and verification by the reference standard was coded and compiled in a comprehensive table (appendix). We assessed the quality of the selected studies using the QUADAS-2 checklist.[19] To define test positivity of the HPV test, we accepted the cutoff proposed by the manufacturer, whereas for cytological tests we considered two cutoffs: atypical squamous cells of undetermined significance (ASC-US) or worse and low-grade squamous intraepithelial lesions (LSIL) or worse.[20]

Statistical analysis

The pooled absolute sensitivity and specificity of the tests were estimated jointly with metandi, a procedure in STATA, based on the bivariate normal model for the logit transforms of sensitivity and specificity, taking the intrinsic correlation between true-positive and false-positive rates and the variability between studies into account.[21,22] We computed the relative sensitivity and specificity of HPV testing on self-samples compared with cytology or HPV testing on clinician-taken samples using metadas, a SAS macro for meta-analysis of diagnostic accuracy studies that allows the inclusion of type of test as a covariate, making comparison of tests possible.[18,23] We also applied this model to assess the effect of study characteristics on the absolute test accuracy. We assessed heterogeneity in the relative sensitivity and specificity (testing on self-sample *vs* clinician-taken samples), in particular the effect of the HPV assays and sampling devices, separately with subgroup meta-analyses and sensitivity analyses.[24,25] We identified influential reports by repeating the meta-analysis and consecutively omitting each individual study.[26] We applied the effective sample-size funnel plot and associated regression test to assess publication bias or small study effects in the meta-analyses of the absolute accuracy.[27] To verify small-sample effects in the relative sensitivity and specificity of HPV testing in self-samples versus clinician-taken samples, we checked asymmetry in funnel plots visually and subsequently tested them statistically by a linear regression based on the efficient score and its variance.[28,29]

Statistical tests were two-sided and statistical significance was defined as p values of less than 0·05. We used STATA (version 10.1) and SAS (version 9.3) for statistical analyses.

Role of the funding source

The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or

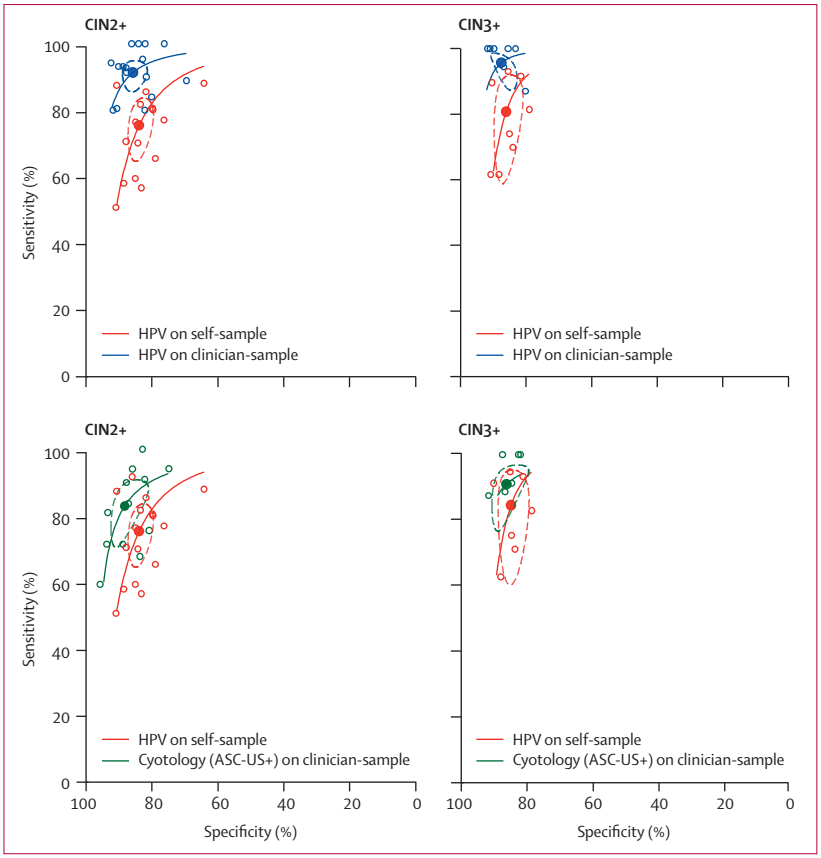


Figure 2: **The accuracy in primary cervical cancer screening, by collection method and grade of cervical intraepithelial neoplasia**

Hollow circles are individual studies, the full curved line is the summary ROC curve, solid circles are the pooled accuracy measures, surrounded by the 95% confidence ellipse (dashed line). CIN2+=cervical intraepithelial neoplasia of grade 2 or worse. CIN3+=cervical intraepithelial neoplasia of grade 3 or worse. HPV=human papillomavirus. ASC-US+=atypical squamous cells of undetermined significance or worse.

	Number of studies	Relative sensitivity (95% CI)	Relative specificity (95% CI)
HPV on self-samples *vs* cytology (ASC-US or worse) on clinician-taken samples			
Outcome of CIN2 or worse	19	0·95 (0·91–0·99)*	0·92 (0·90–0·94)*
Outcome of CIN3 or worse	6	0·99 (0·94–1·06)	0·98 (0·97–0·99)*
HPV on self-samples *vs* cytology (LSIL or worse) on clinician-taken samples			
Outcome of CIN2 or worse	11	1·14 (1·07–1·21)*	0·88 (0·86–0·90)*
Outcome of CIN3 or worse	6	1·19 (1·09–1·29)*	0·90 (0·87–0·94)*
HPV on self-samples *vs* HPV on clinician-taken samples			
Outcome of CIN2 or worse	34	0·88 (0·85–0·91)*	0·96 (0·95–0·97)*
Outcome of CIN3 or worse	12	0·89 (0·83–0·96)*	0·96 (0·93–0·99)*

ASC-US=atypical squamous cells of undetermined significance. CIN2=cervical intraepithelial neoplasia of grade 2. CIN3=cervical intraepithelial neoplasia of grade 3. HPV=assay that picks up high-risk types of the human papillomavirus. *Statistically significant (p<0·05).

Table 3: **Pooled relative sensitivity and specificity of human papillomavirus testing on self-samples versus cytology or human papillomavirus testing on clinician-taken samples, by grade of cervical intraepithelial neoplasia**

writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

We identified 884 articles, of which 34 were papers assessing the clinical accuracy of HPV DNA or RNA testing in self-samples (figure 1, appendix).[15,16,30–61] Because one paper contained three separate studies,[55] data from 36 studies were used in the meta-analysis, altogether including 154 556 women. In all but two selected studies,[47,50] the comparator test was HPV testing on a clinician-taken sample, whereas in 20 reports, the clinician-taken samples were examined cytologically.[15,16,32–34,36,37,39,42–44,47,48,50,51,53,55,61] We requested and obtained unreported accuracy data for the disease threshold CIN2 or worse from three studies.[51,55,61] 16 studies (from 14 papers) were in primary screening of generally healthy women,[33,34,36,39,40,42,43,47,50,51,53,55,58,61] whereas three studies screened high-risk populations.[35,44,45] In 17 reports, women referred to colposcopy were enrolled.[15,16,30–32,37,38,41,46,48,49,52,54,56,57,59,60] One study recruited women under follow-up and women who were treated for cervical precancer.[48] An overview of design, population, and test characteristics and an overview of HPV assays and sampling devices used, which are grouped into five categories (brush, lavage, spatula, swab, and tampon) are available in the appendix.

The methodological quality of the 36 included studies was, overall, moderate to good with average negative scores for the 13 QUADAS items varying between 0% and 31% (4/13), equivocal scores varying between 0% and 58% (7/13), and positive scores varying between 17% (2/13) and 100% (table 1). Risk of bias regarding enrolment of patients was low in 19 (53%), moderate in 16 (44%), and high in one (3%) of 36 studies. Reporting and execution of tests (description of cut-off or blinding of the index test towards comparator and reference test) was adequate in 26 (72%) studies and unclear in ten (28%) studies, but the risk of bias was never assessed as high. The quality of the verification with a reference standard (acceptable validity, blinding towards tests, avoidance of incorporating test results in final conclusion of disease outcome) was good in 32 (89%), moderate in three (8%), and possibly problematic in one (3%) of the 36 studies. The delay between self-sampling, clinician-sampling, and verification with the reference standard was short (6 months or less) in 25 (69%), unreported in nine (25%), and long in two (6%) studies. Partial verification was avoided in 28 (78%) studies but clearly present in eight (22%), whereas differential verification was absent in all but one study.[54] Withdrawal of patients was explained appropriately in 25 (69%) but not in nine (25%) studies. In most studies, uninterpretable results were poorly reported (20 [56%] studies for the assessed tests and 22 [61%] studies for the reference standard).

Table 2 shows the pooled absolute sensitivity and specificity of HPV DNA testing on self-samples and clinician-taken samples and of cytology at cutoff ASC-US or worse and LSIL or worse on clinician-taken samples for detecting underlying CIN2 or worse and CIN3 or worse for each setting separately. Because of the substantial and

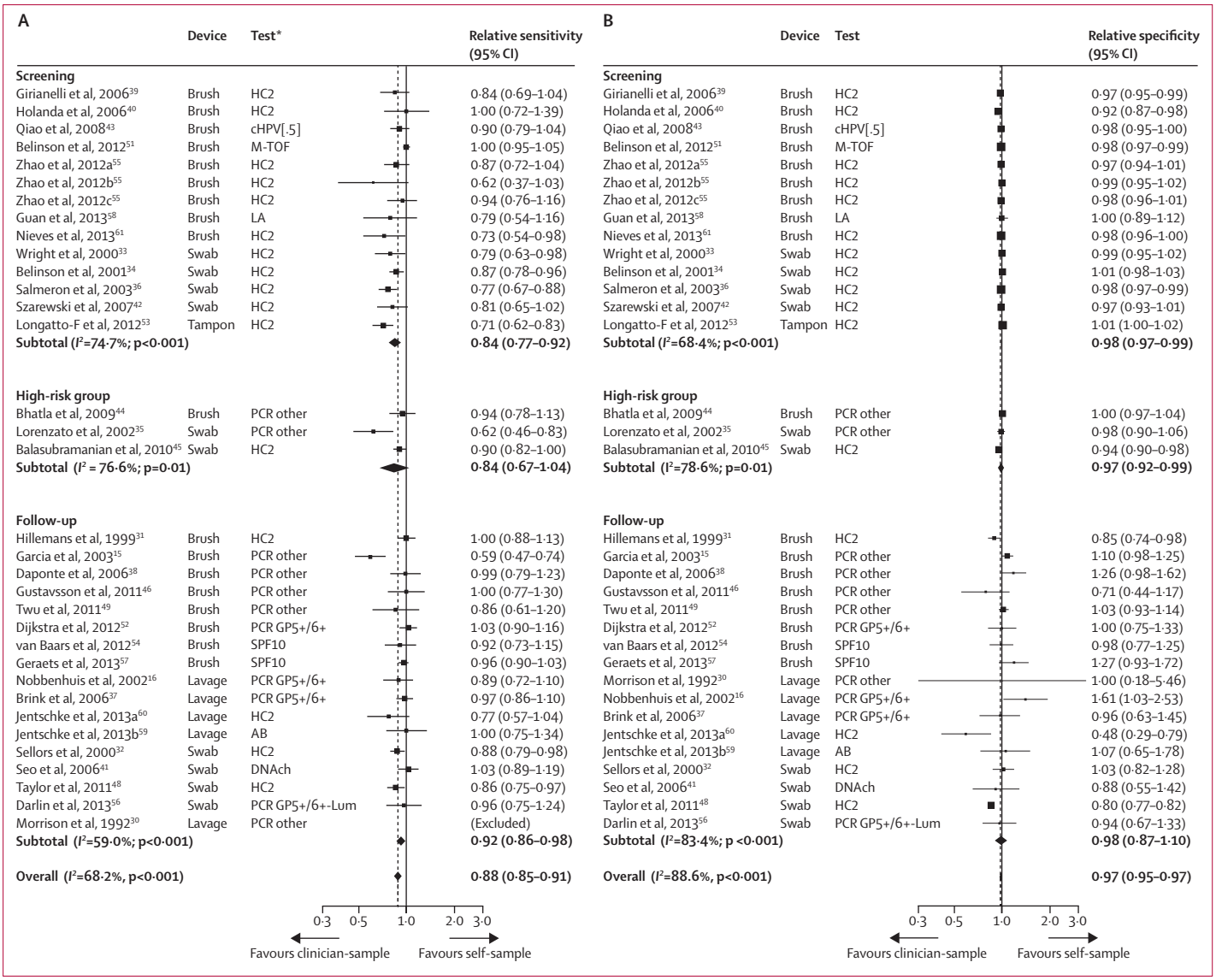


Figure 3: **Relative sensitivity (A) and specificity (B) of human papillomavirus on self-samples versus clinician-taken samples, by clinical setting, for outcome CIN2 or worse**
*See appendix for full test names.

significant variation by setting (p<0·0001; data not shown), we computed no overall pooled accuracy values. In primary screening, the sensitivity of HPV testing on self-samples ranged between 51%[61] and 93%[51] for CIN2 or worse and between 63%[61] and 94%[51] for CIN3 or worse. The pooled sensitivity estimates were 76% (95% CI 69–82) for CIN2 or worse and 84% (72–92) for CIN3 or worse (figure 2, table 2). The specificity for excluding CIN2 or worse ranged between 67%[40] and 93%.[55] HPV testing on clinician-taken samples showed high pooled sensitivity estimates for both CIN2 or worse and CIN3 or worse, and the pooled specificity for excluding CIN2 or worse was 88% (85–91). The sensitivity of cytological testing at cut-off ASC-US or worse on clinician-taken samples was between that of HPV testing in self-samples and clinician-taken samples (table 2). Cytology at the cutoff LSIL or worse showed the lowest pooled sensitivity (for CIN2 or worse), but the highest pooled specificity (for excluding CIN2 or worse; table 2).

The accuracy estimates for CIN2 or worse of HPV testing in self-samples and clinician-taken samples pooled from four studies that included high-risk women were similar to those in primary screening (table 2). However, the accuracy of cytology and of HPV testing for CIN3 or worse varied from the pooled estimates in primary screening, but data in the high-risk testing group were based on only one study.

The pooled sensitivities for CIN2 or worse derived from follow-up studies were not substantially different from those seen in primary screening. HPV testing of

	Number of studies (number of test-device combinations)	Relative sensitivity	Relative specificity
Test			
HC2	18	0·85 (0·81–0·90)*	0·96 (0·93–0·98)*
PCR GP5+/6+	5	0·95 (0·89–1·01)	1·11 (0·95–1·29)
CareHPV (at RLU≥0.5)	1	0·90 (0·79–1·04)	0·98 (0·95–1·00)
CareHPV (at RLU≥1)	1	0·86 (0·73–1·03)	1·00 (0·98–1·02)
PCR-SPF10	2	0·96 (0·89–1·02)	1·10 (0·85–1·41)
Abbott Real Time hrHPV Test	1	1·00 (0·75–1·34)	1·07 (0·65–1·78)
Cervista	1	0·76 (0·70–0·83)*	0·95 (0·94–0·96)*
APTIMA	1	0·64 (0·46–0·90)*	0·99 (0·98–1·01)
DNAchip	1	1·03 (0·89–1·19)	0·88 (0·55–1·42)
Modified GP5+/6+ PCR with Luminex reading	1	0·96 (0·75–1·24)	0·94 (0·67–1·33)
Linear Array	1	0·79 (0·54–1·16)	1·00 (0·89–1·12)
MALDI-TOF	1	1·00 (0·95–1·05)	0·98 (0·97–0·99)*
Other nonGP5+/6+ PCR	7	0·82 (0·66–1·01)	1·02 (0·97–1·07)
Collection device for self-sampling			
Brush	18 (24)	0·89 (0·83–0·94)*	0·98 (0·97–0·99)*
Lavage	5 (6)	0·94 (0·85–1·03)	0·95 (0·68–1·34)
Swab	10	0·86 (0·80–0·92)*	0·95 (0·90–1·01)
Tampon	1	0·71 (0·62–0·83)*	1·01 (1·00–1·02)*
Income status of country			
Low-income and middle-income countries	19 (23)	0·85 (0·79–0·91)*	0·97 (0·95–0·99)*
High-income countries	14 (17)	0·94 (0·90–0·97)*	0·99 (0·93–1·06)

RLU=relative light units. *Statistically significantly different from unity (1 excluded from 95% CI).

Table 4: **Subgroup meta-analysis of the relative sensitivity and specificity of human papillomavirus (self-sampled *vs* clinician sampled) testing for cervical intraepithelial neoplasia of grade 2 or worse, by covariate**

self-samples showed a higher pooled estimate (84% [78–89]), but the difference was not significant (p=0·23). However, the pooled specificity in follow-up settings was substantially and significantly (p<0·0001) lower than the specificity in screening (table 2).

Table 3 shows the pooled relative accuracy measures of HPV testing on self-samples versus the comparator tests on clinician-taken samples. From studies reporting results of multiple tests, we included the combination that yielded the highest relative sensitivity for CIN2 or worse in the meta-analysis under the condition that for a given study the HPV assay always had to be the same in self-samples and clinician-taken samples. The variation of the relative sensitivity and specificity of HPV testing on self-samples versus clinician-taken samples for detection of CIN2 or worse of all the included studies, grouped by clinical setting, is shown in figure 3. We identified substantial heterogeneity in all studies (p<0·001 for all studies). However, we detected no significant heterogeneity in relative sensitivity by clinical setting (p=0·345; data not shown). The variation in relative specificity by setting was also small but statistically significant (p=0·001; data not shown) due to large numbers. Because of the restricted inter-setting variation in the relative accuracy, we did meta-analytical pooling across all clinical settings combined.

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 (table 3, appendix). 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).

HPV testing on self-samples was more sensitive but less specific than cytology using LSIL or worse as a cut-off for CIN2 or worse and CIN3 or worse (table 3, appendix).

HPV testing on self-samples was, on average, less sensitive than HPV testing on clinician-taken samples, for both CIN2 or worse (ratio 0·88) and CIN3 or worse (ratio 0·89; table 3, appendix). Additionally, on self-samples, HPV testing was less specific than on clinician-taken samples in excluding CIN2 or worse (table 3, figure 3).

The variation in the absolute sensitivity and specificity of HPV testing on self-samples to detect CIN2 or worse in primary screening, according to study covariates is shown in the appendix. We saw a statistically significantly higher sensitivity when PCR with MALDI-TOF was used to identify HPV in self-samples than when the Hybrid Capture 2 (HC2) assay (ratio sensitivity of HC2 in self-samples/sensitivity of MALDI-TOF in self-samples for CIN2 or worse, 1·25 [95% CI 1·12–1·39]). The APTIMA test showed a significantly higher specificity than HC2 (ratio 1·04 [1·02–1·07]). Issues of study design or quality of reporting did not alter the clinical accuracy at the exception of one QUADAS item: when the delay between tests was not clearly reported, the sensitivity was significantly lower than when it was clearly reported. We detected no evidence of an effect of involvement of manufacturers of tests or sampling devices on the accuracy estimates.

HPV testing on self-samples was significantly less sensitive and also less specific than such testing on clinician-taken samples to detect CIN2 or worse when HC2 or Cervista were used (table 4). APTIMA testing was less sensitive but not less specific in self-samples versus clinician-taken samples. The sensitivity was similar in both types of samples when testing with GP5+/6+ PCR, SPF10 PCR, Abbott Real Time hrHPV test, DNAchip, modified GP5+/6+ PCR with Luminex reading, or MALDI-TOF. HPV testing on self-samples was less specific than testing on clinician-taken samples when HC2, Cervista, or MALDI-TOF were used. Covariate analyses for sampling device, type of reference test, or completeness of verification generally showed lower sensitivity of HPV testing on self-samples versus such testing on clinician-taken samples (table 4). This finding can be explained by the use of HC2 or other less sensitive HPV assays in most studies (table 5 for device, data not shown for other covariates).

Table 5 shows relative sensitivity and specificity values for tests clinically validated for primary cervical cancer screening on clinician-collected cervical scrapes.[7,62] When

HC2 was used, self-sampling with any device showed lower sensitivity with differences being statistically significant for brushes, swabs, and tampons. 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 (table 5). Also, in the only available study,[59] where the Abbott RT HPV PCR, a GP5+/6+ derivative, was applied, no difference was seen between testing on a self-collected lavage sample and a clinician-collected brush sample.

No differences by country status (high-income *vs* low-income or middle-income) could be discerned in the relative sensitivity. However, the relative specificity was significantly lower than unity in middle-income or low-income countries but not in the only screening study done in a high-income country (table 4).[42]

The effective size funnel plot and associated regression test did not show statistically significant small-study effects in the absolute accuracy of HPV testing on self-samples, or in that of any of the comparator tests on clinician-taken samples (appendix).

We detected no statistically significant asymmetry in the relative specificity of HPV testing in self-sample versus comparator tests (p always >0·10; appendix). However, we detected a statistically significant small-study effect in the relative sensitivity of HPV testing on self-samples versus HPV testing (for CIN2 or worse and CIN3 or worse) or cytology at cut-off ASC-US (for CIN3 or worse) on clinician-taken samples (appendix).

	Number of studies	Relative sensitivity	Relative specificity
HC2			
Brush[31,39,40,44,55,61]	8	0·89 (0·82–0·98)†	0·97 (0·96–0·99)†
Lavage[59,60]	2	0·82 (0·65–1·02)	0·68 (0·35–1·33)
Swab[32-34,36,42,45,48]	7	0·82 (0·86–0·90)†	0·95 (0·89–1·01)
Tampon[53]	1	0·71 (0·62–0·83)†	1·01 (1·00–1·02)†
GP5+/6+ PCR			
Brush[52,54,57]	3	0·95 (0·86–1·04)	1·08 (0·93–1·25)
Lavage[16,37]	2	0·95 (0·85–1·06)	1·23 (0·74–2·05)
Abbott Real Time hrHPV Test			
Lavage[59]	1	1·00 (0·75–1·34)	1·07 (0·65–1·78)

*Restricted to three tests clinically validated for primary cervical cancer screening using samples taken by a clinician.[7] †Statistically significantly different from unity (95% CI excludes 1).

Table 5: **Subgroup meta-analyses of the relative sensitivity and specificity of human papillomavirus (self-sampled *vs* clinician sampled) testing* for cervical intraepithelial neoplasia of grade 2 or worse, by self-collection device**

Discussion

Our findings suggest that screening with an HPV test on self-sampled material can detect, on average, 76% of CIN2 or worse and 84% of CIN3 or worse. The pooled specificity to exclude CIN2 or worse was estimated at 86%. Because the absolute accuracy varied by clinical setting, these values include only primary screening studies. However, the variation of the relative sensitivity and specificity of HPV testing on self-samples compared with tests on clinician-taken samples was low across settings. Such low variation enabled us to pool screening, testing in high-risk populations, and follow-up studies in the meta-analyses of the relative accuracy. The pooled sensitivity of HPV testing on self-samples was statistically significantly lower than HPV testing on a clinician-taken sample. Furthermore, HPV testing on self-samples was also 4% less specific for exclusion of CIN2 or worse. Compared with cytological processing of a cervical cell specimen collected by a health professional and using ASC-US as a cutoff, HPV testing on self-samples was slightly less sensitive and clearly less specific than testing on clinician-taken samples. However, when the threshold for positive cytology was LSIL, HPV testing on a self-sample was 14% more sensitive in detecting CIN2 or worse and 19% more sensitive in detecting CIN3 or worse, but was 12% less specific in excluding CIN2 or worse.

We could discern no obvious collection device effects in the relative sensitivity of HPV testing on self-samples versus clinician-taken samples. Because the included studies did not address comparisons between different sampling devices, and because of the variability in study design and settings, we can draw no strong conclusions about the influence of devices for self-sampling on the accuracy estimates. However, our findings did show obvious test effects. Sensitivity for HPV testing on self-samples was lower than HPV testing on clinician-taken samples when HC2, APTIMA, or Cervista were used. For HC2, this lower sensitivity was consistent for all types of devices separately and could be pooled from 18 studies, whereas findings for Cervista and APTIMA were based on findings from only single studies. Moreover, we saw lower specificity on self-samples for HC2 and Cervista but not for APTIMA. PCR amplification with the GP5+/6+ primers (in five studies[16,37,52,54,57]) and the Abbott RT hrHPV Test (in one study[59]) showed similar sensitivity and specificity on self-samples versus clinician-taken samples. The MALDI-TOF assay was used in only one study,[51] but that study enrolled more than 8000 women attending cervical cancer screening among whom 141 women with CIN3 or worse and 233 women with CIN2 or worse were identified. In this large study, the investigators could assess non-inferiority of MALDI-TOF testing on self-tests versus MALDI-TOF testing on clinician samples, with high precision: relative sensitivity of 1·00 for CIN2 or worse and 1·01 for CIN3 or worse with a lower 95% CI bound greater than 0·90 (p for non-inferiority of <0·0001). MALDI-TOF testing on self-samples showed a small but statistically significant loss in specificity (ratio of 0·98). Some other PCR-based HPV assays also showed a sensitivity which was not statistically significantly lower in self-samples but the precision of equivalency was low (unity included in the 95% CI but lower bound <0·90).

Overall, HPV testing on self-samples was equally sensitive compared with cytology at cutoff ASC-US to detect CIN3 or worse but less sensitive for CIN2 or worse. However, this lower pooled sensitivity for CIN2 or worse was driven by one study with very low sensitivity of GPMY09/11 PCR amplification (49%) on self-samples taken with an endocervical brush in three colposcopy clinics in the USA, Peru, and Mexico.[15] Omission of these outlying findings resulted in a relative sensitivity of 0·99 (95% CI 0·96–1·03) and a relative specificity of 0·97 (0·94–1·00). When positive cytology on clinician-taken samples was defined as LSIL or worse, HPV testing on self-samples was more sensitive in the detection of CIN2 or worse and CIN3 or worse. However, this finding was affected by the outlying high relative sensitivity noted in a large community-based randomised trial,[47] done in Mexico, which compared HC2 on self-samples taken with a conical brush with a conventional Pap smear prepared by a health professional (relative sensitivity of 3·41 for CIN2 or worse and 3·20 for CIN3 or worse).

Although we often judged study design and quality to be non-optimal or insufficiently documented, QUADAS items did not explain the heterogeneity seen in the absolute or relative accuracy estimates (appendix).

In meta-analyses assessing test performance, the pooled parameters are sensitivity and specificity rather than predictive values, which are strongly affected by the background prevalence of disease.[63] However, for a given region or country, compared with sensitivity and specificity values, predictive values are more informative for clinical practice and decision making. We have therefore applied the pooled absolute accuracy estimates of tests on three screening situations with low (0·25%), medium (0·5%), and high (2%; seen in some high-income countries) prevalence of CIN3 or worse to compute post-test probabilities of disease (appendix), in which we regard a risk lower than 1% after a negative test as reassuring, whereas a risk greater than 10% after a positive test would indicate a referral as being appropriate. A negative screening test always yielded a low post-test risk (<1%) of underlying CIN3 or worse but the future risk, considered over a period of 5 years, exceeded 1% after a negative HC2 test on self-sample and after a negative cytology result on a clinician sample in high-risk populations. The cross-sectional positive predictive values on self-samples and clinician-taken samples were usually lower than 10% in low-risk and medium-risk populations, indicating the need for a triage test before referral to colposcopy for diagnostic work-up or treatment. In high-risk populations, the post-test probability of CIN3 or worse was higher than 10% when using cytology or hrHPV testing on clinician samples on a self-sample and also for MALDI-TOF, but not for HC2, on a self-sample.

In terms of the biological plausibility of our findings, the lower clinical sensitivity of HPV testing with HC2, Cervista, or APTIMA on self-samples can be explained by lower loads of high-risk HPV DNA from virus particles in the vagina that are beneath the detection threshold of respective assays, but which can be detected by more analytically sensitive PCR tests.[51,64] Additionally, the lower specificity of HC2, Cervista, and MALDI-TOF might be partly attributable to the increased presence of cross-reacting low-risk HPV types in the vagina but also to presence of high-risk HPV particles in the vagina that have not caused CIN2 or worse.[64]

Our meta-analysis included 36 studies that, together, enrolled more than 150 000 women, potentially allowing for robust estimates with use of statistical procedures recommended by the Cochrane Collaboration. However, only 14 studies took place in a typical screening situation, where self-sampling is of interest. Nonetheless, the relative accuracy comparing testing in self-samples versus clinician-taken samples did not vary substantially by clinical setting, which justified pooling all the studies, and which enabled analysis of the effect of multiple sources of heterogeneity. However, the assessment of the effect of covariates was restricted by paucity of detail in reported aggregated data and sometimes yielded findings on categories based on few or only one study.

Our meta-analysis assessed test accuracies but did not assess the effect of self-sampling on the incidence of cervical cancer (or precancer) in a population. However, if HPV testing on a specimen collected by a woman is analysed with an assay showing similar accuracy on self-samples and clinician samples, it would be likely to result also in reduced incidence of cervical cancer as already shown in randomised trials in which samples were taken by clinicians.[7] The use of an imperfect gold standard based on colposcopy and colposcopy-targeted biopsies could have generated bias. Nonetheless, our main conclusion that HPV testing on self-samples samples was less sensitive than HPV testing on clinician-taken samples was seen for all types of reference testing, including taking random cervical biopsies.

Our findings draw attention to the need for well-designed studies to assess the accuracy of combinations of HPV assays and self-sampling devices. A colposcopy clinic where both a self-sample and clinician-taken sample are obtained, followed by verification by colposcopy and biopsies, offers a valid design in which the absolute sensitivity, as well as the relative sensitivity and specificity, can be assessed without bias on a rather small study sample and which will be also relevant for a primary screening setting. The absolute specificity assessed in a colposcopy clinic is irrelevant for primary screening. Population screening trials can address questions of screening coverage, follow-up compliance of screen-positive women, and positive predictive value for high-grade CIN, but might not enable assessment of absolute specificity. However, the approximate absolute specificity can be assessed without any bias and without need for verifying women with negative screening tests.[65] Particular attention should be given to the assessment of HPV testing on self-samples in low-income and middle

income countries that do not have the infrastructure for cytopathological assessment and have a substantial burden of cervical cancer—in such countries, self-sampling could enable good screening coverage. Longitudinal follow-up studies are also needed. Experts and methodologists should work out well-designed protocols for further studies and define minimal accuracy thresholds that have to be fulfilled to accept a given combination of test and self-collection device in screening as done for screening tests on clinician-taken samples.[62]

Findings from randomised trials published over the past 6 years have shown that HPV screening followed by cytology triage results in a lower incidence of cancer than with cytology screening.[7,66] Because cytology triage is not possible on self-samples, an adequate molecular reflex test should be developed to triage women with HPV. Hyper-methylation of some viral or human genes involved in carcinogenesis have shown promising accuracy profiles and could be applicable on self-samples, but needs further validation.[67,68]

In a screening programme using an HPV assay based on signal-amplification or RNA detection, samples taken by a medical professional rather than self-samples should be chosen, given the superior clinical sensitivity and specificity on clinician-taken samples. In a cytology-based or HPV-based screening programme, HPV-testing on a self-sample can be suggested as an additional strategy to reach women not participating in the regular screening programme. HPV testing on self-samples can also be considered in areas lacking high-quality cytopathology laboratories and where self-sampling could achieve good attendance. Before making a decision on whether to introduce strategies that use self-sampling, adequate transport of samples and communication—with women with a positive self-test being referred to further management—should be assured. In the future, some HPV tests that amplify viral DNA sequences from self-samples might reach similar accuracy as from clinician samples. Such tests might be preferred if validated, available, affordable, and feasible. However, before deciding on HPV screening using self-samples instead of clinician-collected samples, a careful pilot study should assess the feasibility, the clinical accuracy of the combination of the proposed test and the self-collection device, as well as the costs, logistics, and population compliance.

Contributors

MA, PJFS, F-HZ, and AA designed the study, formulated the clinical question, and identified the PICOS components. SM, CB, RB, FV, and MA did the literature searches. MA designed the data extraction form. MA, ES, PS, VMJV, LD, and FV extracted the data. MA and FV did the statistical analyses. MA wrote the paper. FV, PS, VMJV, ES, LD, SM, CB, RB, F-HZ, PH, and AA critically revised subsequent drafts.

Conflicts of interest

PJFS has received speaker's fees from Roche, Abbott, Qiagen, and Gen-Probe, and is shareholder of Self-screen. PH received research grants or lecture fees from Abbott, Hologic, and Roche. MA, FV, VMJV, ES, LD, SM, CB, RB, F-HZ, and AA declare that they have no conflicts of interest.

Acknowledgments

This study was funded by the seventh Framework Program of DG Research of the European Commission (PREHDICT Project [grant number 242061] and CoheaHr Project [603019], coordinated by the VU University Medical Center of Amsterdam, the Netherlands), the Belgian Foundation against Cancer (Brussels, Belgium), the International Agency for Research on Cancer (IARC, Lyon, France), and the German Guideline Program in Oncology (Hannover, Germany). We thank J L Belinson and F-H Zhao for the contribution of additional data and V Liausson for high-quality eps files for figures 2 and 3.

References

1 Cogliano V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, and the WHO International Agency for Research on Cancer. Carcinogenicity of human papillomaviruses. *Lancet Oncol* 2005; **6:** 204.

2 IARC Monograph Working Group. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol 90: Human Papillomaviruses. Lyon: IARC Press, 2007.

3 Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J. Chapter 9: clinical applications of HPV testing: a summary of meta-analyses. *Vaccine* 2006; **24** (suppl 3)**:** S3, 78–89.

4 European Commission. European Guidelines for Quality Assurance in Cervical Cancer Screening. Luxembourg: Office for Official Publications of the European Communities; 2008.

5 Wright TC Jr, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D, and the 2006 American Society for Colposcopy and Cervical Pathology-sponsored Consensus Conference. 2006 consensus guidelines for the management of women with cervical intraepithelial neoplasia or adenocarcinoma in situ. *Am J Obstet Gynecol* 2007; **197:** 340–45.

6 Jordan J, Arbyn M, Martin-Hirsch P, et al. European guidelines for quality assurance in cervical cancer screening: recommendations for clinical management of abnormal cervical cytology, part 1. *Cytopathology* 2008; **19:** 342–54.

7 Arbyn M, Ronco G, Anttila A, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine* 2012; **30** (suppl 5)**:** F88–999.

8 Saslow D, Solomon D, Lawson HW, et al, and the ACS-ASCCP-ASCP Cervical Cancer Guideline Committee. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin* 2012; **62:** 147–72.

9 Arbyn M, de Sanjosé S, Saraiya M, et al. EUROGIN 2011 roadmap on prevention and treatment of HPV-related disease. *Int J Cancer* 2012; **131:** 1969–82.

10 Snijders PJ, Verhoef VM, Arbyn M, 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* 2013; **132:** 2223–36.

11 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; **104:** e159–66.

12 Ogilvie GS, Patrick DM, Schulzer M, et al. Diagnostic accuracy of self collected vaginal specimens for human papillomavirus compared to clinician collected human papillomavirus specimens: a meta-analysis. *Sex Transm Infect* 2005; **81:** 207–12.

13 Petignat P, Faltin DL, Bruchim I, Tramèr MR, Franco EL, Coutlée F. Are self-collected samples comparable to physician-collected cervical specimens for human papillomavirus DNA testing? A systematic review and meta-analysis. *Gynecol Oncol* 2007; **105:** 530–35.

14 Stewart DE, Gagliardi A, Johnston M, et al, and the HPV Self-collection Guidelines Panel. Self-collected samples for testing of oncogenic human papillomavirus: a systematic review. *J Obstet Gynaecol Can* 2007; **29:** 817–28.

15 Garcia F, Barker B, Santos C, et al. Cross-sectional study of patient- and physician-collected cervical cytology and human papillomavirus. *Obstet Gynecol* 2003; **102:** 266–72.

16 Nobbenhuis MA, Helmerhorst TJ, van den Brule AJ, et al. Primary screening for high risk HPV by home obtained cervicovaginal lavage is an alternative screening tool for unscreened women. *J Clin Pathol* 2002; **55:** 435–39.

17 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151:** 264–69.

18 Diagnostic Test Accuracy Working Group. Handbook for Diagnostic Test Accuracy Reviews 2011. http://srdta.cochrane.org/handbook-dta-reviews (accessed Dec 1, 2013).

19 Whiting PF, Rutjes AW, Westwood ME, et al, and the QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155:** 529–36.

20 Solomon D, Davey D, Kurman R, et al, and the Forum Group Members, and the Bethesda 2001 Workshop. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002; **287:** 2114–19.

21 Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007; **8:** 239–51.

22 Harbord RM, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J* 2009; **9:** 211–29.

23 Takwoingi Y. Diagnostic Test AccuracyWorking Group. METADAS: A SAS macro for meta-analysis of diagnostic accuracy studies. http://srdta.cochrane.org/sites/srdta.cochrane.org/files/uploads/MetaDAS%20Quick%20Reference%20v1.3%20May%202012.pdf (accessed Dec 1, 2013).

24 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7:** 177–88.

25 Harris R, Bradburm M, Deeks J, et al. Metan: fixed- and random-effects meta-analysis. *Stata J* 2008; **8:** 3–26.

26 Tobias A. Assessing the influence of a single study in the meta-analysis estimate. *Stata Tech Bull* 1998; **8:** 108–10.

27 Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005; **58:** 882–93.

28 Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med* 2006; **25:** 3443–57.

29 Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. *Stata J* 2009; **9:** 197–210.

30 Morrison EAB, Goldberg GL, Hagan RJ, Kadish AS, Burk RD. Self-administered home cervicovaginal lavage: a novel tool for the clinical-epidemiologic investigation of genital human papillomavirus infections. *Am J Obstet Gynecol* 1992; **167:** 104–07.

31 Hillemanns P, Kimmig R, Hüttemann U, Dannecker C, Thaler CJ. Screening for cervical neoplasia by self-assessment for human papillomavirus DNA. *Lancet* 1999; **354:** 1970.

32 Sellors JW, Lorincz AT, Mahony JB, et al. Comparison of self-collected vaginal, vulvar and urine samples with physician-collected cervical samples for human papillomavirus testing to detect high-grade squamous intraepithelial lesions. *CMAJ* 2000; **163:** 513–18.

33 Wright TC Jr, Denny L, Kuhn L, Pollack A, Lorincz A. HPV DNA testing of self-collected vaginal samples compared with cytologic screening to detect cervical cancer. *JAMA* 2000; **283:** 81–86.

34 Belinson J, Qiao YL, Pretorius R, et al. Shanxi Province Cervical Cancer Screening Study: a cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. *Gynecol Oncol* 2001; **83:** 439–44.

35 Lorenzato FR, Singer A, Ho L, et al. Human papillomavirus detection for cervical cancer prevention with polymerase chain reaction in self-collected samples. *Am J Obstet Gynecol* 2002; **186:** 962–68.

36 Salmerón J, Lazcano-Ponce E, Lorincz A, et al. Comparison of HPV-based assays with Papanicolaou smears for cervical cancer screening in Morelos State, Mexico. *Cancer Causes Control* 2003; **14:** 505–12.

37 Brink AA, Meijer CJ, Wiegerinck MA, et al. High concordance of results of testing for human papillomavirus in cervicovaginal samples collected by two methods, with comparison of a novel self-sampling device to a conventional endocervical brush. *J Clin Microbiol* 2006; **44:** 2518–23.

38 Daponte A, Pournaras S, Mademtzis I, et al. Evaluation of HPV 16 PCR detection in self- compared with clinician-collected samples in women referred for colposcopy. *Gynecol Oncol* 2006; **103:** 463–66.

39 Girianelli VR, Thuler LC, Szklo M, et al. Comparison of human papillomavirus DNA tests, liquid-based cytology and conventional cytology for the early detection of cervix uteri cancer. *Eur J Cancer Prev* 2006; **15:** 504–10.

40 Holanda F Jr, Castelo A, Veras TM, de Almeida FM, Lins MZ, Dores GB. Primary screening for cervical cancer through self sampling. *Int J Gynaecol Obstet* 2006; **95:** 179–84.

41 Seo SS, Song YS, Kim JW, Park NH, Kang SB, Lee HP. Good correlation of HPV DNA test between self-collected vaginal and clinician-collected cervical samples by the oligonucleotide microarray. *Gynecol Oncol* 2006; **102:** 67–73.

42 Szarewski A, Cadman L, Mallett S, et al. Human papillomavirus testing by self-sampling: assessment of accuracy in an unsupervised clinical setting. *J Med Screen* 2007; **14:** 34–42.

43 Qiao YL, Sellors JW, Eder PS, et al. A new HPV-DNA test for cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. *Lancet Oncol* 2008; **9:** 929–36.

44 Bhatla N, Dar L, Patro AR, et al. Can human papillomavirus DNA testing of self-collected vaginal samples compare with physician-collected cervical samples and cytology for cervical cancer screening in developing countries? *Cancer Epidemiol* 2009; **33:** 446–50.

45 Balasubramanian A, Kulasingam SL, Baer A, et al. Accuracy and cost-effectiveness of cervical cancer screening by high-risk human papillomavirus DNA testing of self-collected vaginal samples. *J Low Genit Tract Dis* 2010; **14:** 185–95.

46 Gustavsson I, Sanner K, Lindell M, et al. Type-specific detection of high-risk human papillomavirus (HPV) in self-sampled cervicovaginal cells applied to FTA elute cartridge. *J Clin Virol* 2011; **51:** 255–58.

47 Lazcano-Ponce E, Lorincz AT, Cruz-Valdez A, et al. Self-collection of vaginal specimens for human papillomavirus testing in cervical cancer prevention (MARCH): a community-based randomised controlled trial. *Lancet* 2011; **378:** 1868–73.

48 Taylor S, Wang C, Wright TC, Denny L, Kuhn L. A comparison of human papillomavirus testing of clinician-collected and self-collected samples during follow-up after screen-and-treat. *Int J Cancer* 2011; **129:** 879–86.

49 Twu NF, Yen MS, Lau HY, Chen YJ, Yu BK, Lin CY. Type-specific human papillomavirus DNA testing with the genotyping array: a comparison of cervical and vaginal sampling. *Eur J Obstet Gynecol Reprod Biol* 2011; **156:** 96–100.

50 Wikström I, Lindell M, Sanner K, Wilander E. Self-sampling and HPV testing or ordinary Pap-smear in women not regularly attending screening: a randomised study. *Br J Cancer* 2011; **105:** 337–39.

51 Belinson JL, Du H, Yang B, et al. Improved sensitivity of vaginal self-collection and high-risk human papillomavirus testing. *Int J Cancer* 2012; **130:** 1855–60.

52 Dijkstra MG, Heideman DA, van Kemenade FJ, et al. Brush-based self-sampling in combination with GP5+/6+-PCR-based hrHPV testing: high concordance with physician-taken cervical scrapes for HPV genotyping and detection of high-grade CIN. *J Clin Virol* 2012; **54:** 147–51.

53 Longatto-Filho A, Naud P, Derchain SF, et al. Performance characteristics of Pap test, VIA, VILI, HR-HPV testing, cervicography, and colposcopy in diagnosis of significant cervical pathology. *Virchows Arch* 2012; **460:** 577–85.

54 van Baars R, Bosgraaf RP, ter Harmsel BW, et al Dry storage and transport of a cervicovaginal self-sample by use of the Evalyn Brush, providing reliable human papillomavirus detection combined with comfort for women. *J Clin Microbiol* 2012; **50:** 3937–43.

55 Zhao FH, Lewkowitz AK, Chen F, et al. Pooled analysis of a self-sampling HPV DNA Test as a cervical cancer primary screening method. *J Natl Cancer Inst* 2012; **104:** 178–88.

56 Darlin L, Borgfeldt C, Forslund O, Hénic E, Dillner J, Kannisto P. Vaginal self-sampling without preservative for human papillomavirus testing shows good sensitivity. *J Clin Virol* 2013; **56:** 52–56.

57 Geraets DT, van Baars R, Alonso I, et al. Clinical evaluation of high-risk HPV detection on self-samples using the indicating FTA-elute solid-carrier cartridge. *J Clin Virol* 2013; **57:** 125–29.

58 Guan Y, Gravitt PE, Howard R, et al. Agreement for HPV genotyping detection between self-collected specimens on a FTA cartridge and clinician-collected specimens. *J Virol Methods* 2013; **189:** 167–71.

59 Jentschke M, Soergel P, Hillemanns P. Evaluation of a multiplex real time PCR assay for the detection of human papillomavirus infections on self-collected cervicovaginal lavage samples. *J Virol Methods* 2013; **193:** 131–34.

60 Jentschke M, Lange V, Soergel P, Hillemanns P. Enzyme-linked immunosorbent assay for p16(INK4a) – a new triage test for the detection of cervical intraepithelial neoplasia? *Acta Obstet Gynecol Scand* 2013; **92:** 160–64.

61 Nieves L, Enerson CL, Belinson S, et al. Primary cervical cancer screening and triage using an mRNA human papillomavirus assay and visual inspection. *Int J Gynecol Cancer* 2013; **23:** 513–18.

62 Meijer CJLM, 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:** 516–20.

63 Diagnostic Test Accuracy Working Group. Handbook for diagnostic test accuracy reviews. Birmingham: Cochrane Collaboration, DTA Working Group, 2013. http://srdta.cochrane.org/handbook-dta-reviews (accessed Jan 2, 2014).

64 Belinson JL, Hu S, Niyazi M, et al. Prevalence of type-specific human papillomavirus in endocervical, upper and lower vaginal, perineal and vaginal self-collected specimens: implications for vaginal self-collection. *Int J Cancer* 2010; **127:** 1151–57.

65 Arbyn M, Ronco G, Cuzick J, Wentzensen N, Castle PE. How to evaluate emerging technologies in cervical cancer screening? *Int J Cancer* 2009; **125:** 2489–96.

66 Ronco G, Giorgi-Rossi P, Carozzi F, et al, and the New Technologies for Cervical Cancer screening (NTCC) Working Group. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010; **11:** 249–57.

67 Snellenberg S, De Strooper LM, Hesselink AT, et al. Development of a multiplex methylation-specific PCR as candidate triage test for women with an HPV-positive cervical scrape. *BMC Cancer* 2012; **12:** 551.

68 Wentzensen N, Sun C, Ghosh A, et al. Methylation of HPV18, HPV31, and HPV45 genomes and cervical intraepithelial neoplasia grade 3. *J Natl Cancer Inst* 2012; **104:** 1738–49.

Cancer Epidemiology 41 (2016) 152–158

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

Surveillance of effects of HPV vaccination in Belgium

Marc Arbyn[a,*], Davy Vanden Broeck[b,c,d], Ina Benoy[c,d], Johannes Bogers[c,d], Christophe Depuydt[c], Marleen Praet[e], Philippe De Sutter[f], Anne Hoorens[g], Esther Hauben[h], Willy Poppe[i], Marc Van Ranst[j], Philippe Delvenne[k], Stephanie Gofflot[k], Michel Pétein[l], Frans Engelen[m], Alain Vanneste[n], Lode Op De Beeck[o], Pierre Van Damme[p], Marleen Temmerman[b,q,r], Steven Weyers[r]

[a] Unit of Cancer Epidemiology/Belgian Cancer Centre, Scientific Institute of Public Health, J. Wytsmanstreet 14, B1050 Brussels, Belgium
[b] International Centre for Reproductive Health (ICRH), Ghent University, Ghent, Belgium
[c] Algemeen Medisch Labo, Sonic Healtcare, Antwerp, Belgium
[d] AMBIOR, Laboratory for Cell Biology & Histology, University of Antwerp, Antwerp, Belgium
[e] N. Goormachtigh Institute for Pathology, Ghent University, Ghent, Belgium
[f] Department of Gynaecology & Oncology, UZ Brussel, Free University of Brussels, Brussels, Belgium
[g] Department of Pathology, VUB, Brussels, Belgium
[h] Department of Pathology, UZ Leuven, Leuven, Belgium
[i] Department of Gynaecology and Obstetrics, UZ Leuven, Leuven, Belgium
[j] Department of Microbiology and Immunology, Rega Institute, KU Leuven, Leuven, Belgium
[k] Department of Pathology, ULg, Liège, Belgium
[l] Institut de Pathologie et de Génétique, Charlerloi, Belgium
[m] Centrum voor Medische Analyse, Herentals, Belgium
[n] Pathological Anatomy, AZ Groeninge, Kortrijk, Belgium
[o] Pathology Mariaziekenhuis Noord Limburg, Overpelt, Belgium
[p] Vaccine & Infectious Disease Institute (VAXINFECTIO), Antwerp University, Antwerp, Belgium
[q] Reproductive Health and Research, World Health Organization, Geneva, Switzerland
[r] Department of Gynaecology and Obstetrics, Ghent University, Ghent, Belgium

ARTICLE INFO

Article history:
Received 31 August 2015
Received in revised form 19 November 2015
Accepted 22 December 2015
Available online 16 February 2016

Keywords:
Cervical cancer
Screening
Surveillance
HPV
Human papillomavirus
HPV vaccination

ABSTRACT

Background: Early effects of HPV (human papillomavirus) vaccination are reflected by changes observable in young women attending cervical cancer screening.
Subject and methods: The SEHIB study included HPV geno-typing of ~6000 continuous and 650 pathological cervical cell specimen as well as biopsies, collected from women in Belgium in 2010–2014. Data were linked to vaccination status.
Results: HPV vaccination offered protection among women aged <30 years against infection with HPV16 (vaccine effectiveness [VE] = 67%, 95% CI: 48–79%), HPV18 (VE = 93%, 95% CI: 52–99%), and high-risk HPV (VE = 16%, 95% CI: 2–29%). Vaccination protected also against cytological lesions. Vaccination protected against histologically confirmed lesions: significantly lower absolute risks of CIN1+ (risk difference [RD] = −1.6%, 95% CI: −2.6% to −0.7%) and CIN3+ associated with HPV16/18 (RD = −0.3%, 95% CI −0.6% to −0.1%). Vaccine effectiveness decreased with age. Protection against HPV16 and 18 infection was significant in all age groups, however no protection was observed against cytological lesions associated with these types in age-group 25–29.
Conclusion: The SEHIB study demonstrates the effectiveness of HPV vaccination in Belgian young women in particular in age group 18–19. Declining effectiveness with increasing age may be explained by higher tendency of women already exposed to infection to get the vaccine.
© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In Belgium, a country in Western Europe of 11 million inhabitants, 623 new cases of cervical cancer were noted in 2011, and an estimated 275 women died from it [1,2]. Up to

* Corresponding author.
E-mail address: marc.arbyn@wiv-isp.be (M. Arbyn).

http://dx.doi.org/10.1016/j.canep.2015.12.011
1877-7821/© 2016 Elsevier Ltd. All rights reserved.

recently, cervical cancer screening was mainly opportunistic, characterised by a moderate coverage, too frequent screening and a poorly documented quality [3,4]. Sixty percent of women aged 25–64 years have had a Pap smear taken in the last three years [5]. Eighty two percent of all interpreted Pap smears are taken from women belonging to the target age group (25–64 years), whereas 10% and 8% of smears are taken from women younger than 25 years or older than 64 years, respectively [5]. In 2013, a new call-recall programme has been set up in the Flemish Region, where all women, without recent cervix cytology record in the cancer registry, are invited in agreement with European guidelines. In the other two regions (Wallonia, Brussels), no organised programme is in place.

In Belgium, HPV vaccines are reimbursed partially, via the National Health Insurance Institute (NHII), for girls in the range 12–18 years. There is still a personal contribution for the parents of 12€ (or 8€ for socially vulnerable categories)/dose, whereas the rest of the vaccine cost (∼110€/dose) is paid by the NHII. An estimate was made of the coverage generated through this opportunistic HPV vaccination from individual reimbursement claims between November 2007 and May 2012. The 3-dose HPV coverage reached over this period showed a decreasing trend by birth cohort (C): 50% in C1992, 46% for C1993, 37% for C1994 [6].

Free school-based vaccination of girls in the first year of secondary education was introduced by the Flemish Community in 2010, using a three-doses schedule of the quadrivalent vaccine. The 3-dose coverage was 81% and 79% for the school years 2010–11 and 2011–12, respectively [7]. In 2011, also the French community introduced free vaccination with the bivalent vaccine targeting girls in the second year of secondary education. The 3-dose coverage in the French Community for the 2012–13 programme was estimated at 29% [8]. In 2014, health authorities of the Flemish and French Communities decided to switch to a two-doses vaccination schedule with the bivalent vaccine.

The SEHIB (Surveillance of Effects of HPV Immunisation in Belgium) study addressed three objectives: 1) to assess the prevalence of HPV infection and of cervical cytological lesions at the onset of HPV vaccination; 2) to assess the HPV vaccination coverage; and 3) to measure the early impact of HPV vaccination on the burden of HPV infection and HPV-related cervical disease using archived cervical cell and tissue material from women attending cervical cancer screening in Belgium between end 2010 and early 2014.

2. Methods

2.1. Participating centres, enrolment of women

The departments of gynaecology and cytopathology of four Belgian universities (Brussels, Ghent, Leuven, Liège) as well as four large regional cytopathology laboratories (Charleroi, Herentals, Kortrijk and Overpelt) participated in SEHIB. The centres were chosen based on yearly work load of Pap smears and geographical location. All centres were requested to enrol at least consecutive 600 women attending cervical cancer screening with the purpose to reach together a sample size of 6000 specimen. Enrolment was restricted to women aged 18–64 years, with at least 300 women younger than 30 per centre.

The gynaecological departments of the university centres recruited only women whose cervical samples were interpreted in the cytopathology laboratory of the same centres and who had given informed consent. The HPV vaccination status was queried during the gynaecological consultation. The coordinates of the physician having administrated the vaccine were recorded during the first year in two university departments. Subsequently the physician was contacted by the SEHIB coordinating centre to validate the reported vaccination status. Vaccination status was defined as positive when anamnesis indicated that the woman had received at least one dose of HPV vaccine. The recruitment of continuous screening samples took place in phase-1. During phase-2, the university centres enrolled additional pathologic cytology samples. The inclusion of specimen for the SEHIB took place between November 2010 and July 2014.

2.2. Cytology collection, sample processing and HPV testing

Cervical cell samples were collected by the woman's gynaecologist or GP and prepared according to current European guidelines [9]. Cytological results were reported following the Bethesda System [10]. Subsequently, residual LBC material of selected specimen was stored in the original vials at 4 °C, before shipping to the AML laboratory (Antwerp, Belgium) for HPV testing with a validated real-time qPCR assay targeting type-specific sequences of the E6 or E7 genes of 17 HPV types (HPV6 E6, HPV11 E6, HPV16 E7, HPV18 E7, HPV31 E6, HPV33 E6, HPV35 E6, HPV39 E7, HPV45 E7, HPV51 E7, HPV52 E7, HPV53 E6, HPV56 E7, HPV58 E7, HPV59 E7, HPV66 E6 and HPV68 E7) and the L1 gene of HPV67, as described previously [11,12].

Biopsies taken at the four universities from SEHIB women referred to colposcopy were also shipped to the AML laboratory for HPV genotyping after reviewing of the original histological diagnosis by an expert pathologist.

2.3. Data management

A pseudonymous SEHIB code was used to link cytology, histology and virology files with the demographic and clinical data (municipality, birth year, reason of testing, HPV vaccination status). Data compilation and statistical analyses were performed at the Scientific Institute of Public Health, Brussels. Approval of SEHIB was received from the Ethical Review Board of the University Hospital of Ghent (EC approval Nr. B67020109075), which was subsequently endorsed by the other centres.

2.4. Statistical analysis

Considerations regarding the computation of the study sample size and statistical power can be found in the supplementary material. Study records were weighted according to the phase and the reason for Pap smear taking by imposing the distribution of cytological findings observed in the screening population in phase-1 to phase-2. The purpose was to generate a virtual mother population that was substantially larger than the actual study size and to maximize power of the study without necessity to perform HPV testing in all subjects of the mother population. The computation of the reweighted population is shown in Supplementary Table 2.

Because of the oversampling of young women, and to account for differences in the age composition between study centres, directly age-standardised prevalences were computed, using the truncated Belgian female 2012 population aged 18–65 years as reference. The HPV vaccination coverage was estimated from the anamnestic data combined with information obtained from the vaccinating physician. The impact of vaccination status on occurrence of cytological and histological abnormalities associated with HPV vaccine types and irrespective of HPV types was assessed by computing VE (vaccine effectiveness) and RD (risk differences), where VE = 1-RR (risk ratio), $RR = P_{vacc}/P_{n.vacc}$, P_{vacc} being the prevalence of infection or lesion in vaccinated and $P_{n.vacc}$ being the prevalence in non-vaccinated women; $RD = P_{vacc} - P_{n.vacc}$. Perfect vaccine protection corresponds with VE = 100%, absence of protection with 0%. Effectiveness is considered as statistically

significant when VE > 0% and the 95% confidence interval excludes zero, with $p < 0.05$. Statistical significance was defined at $p < 0.05$. Statistical analyses were performed with Stata, version 13 (StataCorp LP, College Station, TX, USA).

3. Results

3.1. Study characteristics,

In total, cervical cell specimen from 6,630 women were included, among whom 5,980 in phase-1 and 650 in phase-2. The geographical origin of enrolled women is shown in Table S3 and Fig. S1. Forty seven percent of samples were from women younger than 30 and 53% were thirty or older (Figure S2). The four regional laboratories provided 45% of the samples (only phase 1), whereas the four university centres contributed 45% and 10%, in phase 1 and 2, respectively (Table S4). The large majority of Pap smears (91%) were taken for screening, the other 9% were collected for other reasons (Table S5). Three percent of cytological samples were judged inadequate. However, after recoding the samples with interpretation for squamous or glandular cells as satisfactory only 0.4% were considered as actually inadequate (Table S6). The age-standardized prevalence of cytological abnormalities in the Belgian screening population was derived from phase-1 (Tables S7 and S8): ASC-US (atypical squamous cells of undetermined significance, 3.7%), LSIL (low-grade squamous intraepithelial lesion, 2.7%), ASC-H (atypical squamous cells where HSIL cannot be excluded, 0.5%), HSIL (high-grade squamous intraepithelial lesion, 0.6%), AGC (atypical glandular cells, 0.1%) and cervical cancer (0.007%). The variation in the prevalence of cytological abnormalities by age is shown in Fig. 1.

3.2. HPV findings, cyto-virological correlation

The proportion of specimen not allowing PCR amplification of *b-globin* was 0.7% (Table S9). In the screening population, the prevalence of hr (high-risk) HPV infection reached a maximum of 20.4% in the age group 20–24 and decreased thereafter up to age group 45–49, interrupted by a peak in age group 50–54 which was followed by a further decrease in the older groups (Fig. 2). The prevalence of the separate HPV types showed similar age patterns with the exception of the high-risk types HPV45, HPV68 and low-risk type HPV11 that showed low rates (<1%) at all ages. Table 1 gives the age-standardized prevalence of infection with hrHPV, with 18 type-specific HPV infections by cytological categories. In general, the prevalences increased by cytological severity of the lesion. The HSIL/NILM (negative for intraepithelial lesion or malignancy) ratio was highest for HPV18 (26.7), HPV58 (23.5) and HPV16 (18.6). HPV16 was the most frequent HPV type in all cytological categories.

3.3. Histological findings

Hundred seventy six biopsies were retrieved from women enrolled in SEHIB. However, 15 were considered as unsatisfactory at quality control review yielding 161 specimen for definite histological interpretation: 38 cases without dysplasia, 68 CIN1 (cervical intraepithelial neoplasia of grade 1), 34 CIN2, 19 CIN3 and 2 cases of squamous cervical cancer (Table S10). The proportion of cytological abnormalities with adequate biopsies varied from 4% for women with ASC-US to 67% for women with cytological suspicion of cancer (Table S11). Eighty three percent of the original histological interpretations (dichotomized as ≤CIN1 and CIN2+) agreed with the definite diagnosis (kappa = 0.66, Table S12). The hrHPV and HPV16 positivity rate in cytological specimen rose with the severity of the histological outcome. Ninety seven percent of the cases of CIN2 and all of CIN3 and cancer were hrHPV positive. HPV16 or HPV18 were present in 63%, 83% and 100% in CIN2, CIN3 and cancer, respectively (Table S13). The virological agreement between cytology and histology specimen was 84% ($\kappa = 0.66$) for HPV16, 95% ($\kappa = 0.57$) for HPV18, 76% ($\kappa = 0.46$) for hrHPV and 67% ($\kappa = 0.51$) for other non-HPV16/18 hrHPV types (Table S14).

3.4. Vaccination status

From November 2010 to October 2011, the vaccination status of women enrolled at the university centres 5 and 8 was verified for those who self-reported that they were vaccinated against HPV. For all women younger than 35 years, whose physician-vaccinator could be contacted, this information was confirmed, whereas for older women, the information was wrong at the exception of one woman aged 35 in phase-2. After October 2011, verification was stopped, accepting the self-reported vaccination status for all women <35 years of age as probable. Accepting this assumption, vaccination coverage could be estimated as: 68% (95% CI: 52–78%) for women younger than 20 years, 33% (95% CI: 29–38%) in age group 20–24 and 9% (95% CI: 7–11%) in age group 25–29, and 0.3% (95% CI: 0–1%) in age group 30–39, whereas zero among older women (Fig. 3). Women with unknown vaccination status were excluded from analysis.

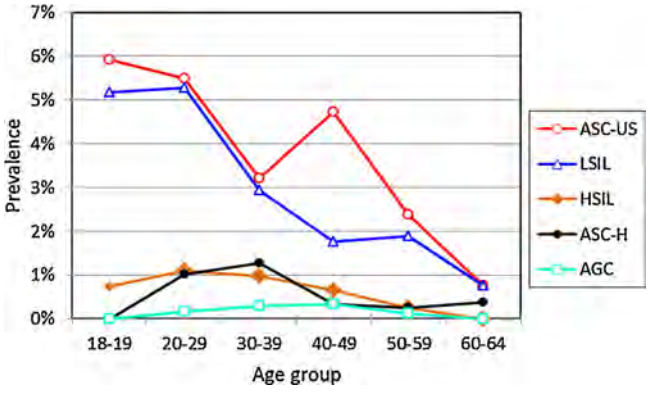


Fig. 1. Prevalence of cervical lesions by cytological category (restricted to the screening population of phase-1), by age group. Abbreviations: ASC-US: atypical squamous cells of undetermined significance, LSIL: low-grade squamous intraepithelial lesion, HSIL: high-grade squamous intraepithelial lesion, ASC-H: atypical squamous cells where HSIL cannot be excluded, AGC: atypical squamous cells.

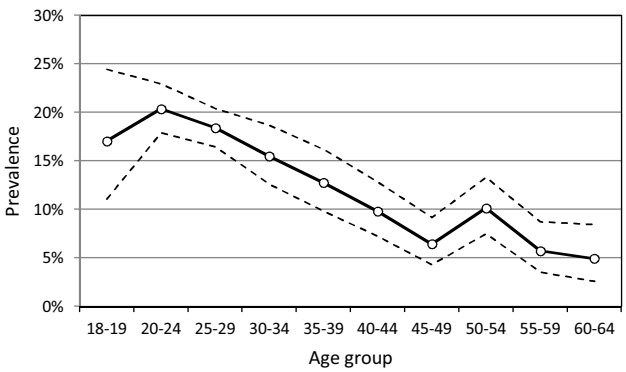


Fig. 2. Prevalence of hrHPV infection (and 95% confidence interval [interrupted lines]), by age group, in the screening population of the SEHIB study.

Table 1
Age-standardised prevalence of cytological findings and of HPV infections by cytological category in the screening population.

Types	NILM	ASC-US	LSIL	ASC-H	HSIL+	Total	Ratio hrHPV in HSIL+/NILM
Prevalence of cytological findings							
N	5096	242	205	34	47	5,624	
%	92.4	3.7	2.7	0.5	0.6	92.4	
Prevalence of HPV infections (in %)							
High-risk types							
hrHPV(13)	9.2	20.0	46.5	12.2	81.9	9.2	8.9
HPV 16	2.5	5.8	15.3	7.3	47.1	3.5	18.6
HPV 18	0.6	1.7	3.3	1.0	16.0	0.9	26.7
HPV 31	1.9	3.2	4.3	1.5	18.4	2.2	9.8
HPV 33	0.4	0.5	1.5	1.5	0.9	0.5	2.4
HPV 35	0.5	0.1	0.7	0.0	7.0	0.5	15.4
HPV 39	1.0	2.7	5.3	0.0	0.0	1.2	0.0
HPV 45	0.3	0.1	0.2	0.0	0.0	0.3	0.0
HPV 51	1.3	4.1	7.3	1.0	9.2	1.7	7.3
HPV 52	0.9	2.0	6.0	1.0	5.0	1.2	5.7
HPV 56	0.8	2.8	8.2	1.0	4.3	1.2	5.1
HPV 58	0.6	2.9	2.8	1.0	14.3	0.9	23.5
HPV 59	0.5	2.3	7.7	2.4	4.3	0.8	8.4
HPV 68	0.1	0.9	0.3	0.0	0.5	0.1	4.9
HPV1618	3.0	7.0	16.2	8.3	56.8	4.2	18.8
hrHPVnon1618	12.2	26.9	62.7	20.5	138.8	15.7	11.3
Intermediate risk types							
HPV 53	0.8	2.5	11.2	3.1	0.5	1.2	0.6
HPV 66	0.7	1.1	8.5	1.7	2.6	1.0	3.4
HPV 67	1.0	2.3	4.1	2.5	0.5	1.2	0.5
Low-risk types causing genital warts							
HPV 06	0.4	0.1	4.0	1.0	0.5	0.5	1.1
HPV 11	0.05	0.1	2.7	0.0	0.0	0.1	0.0

HPV: human papillomavirus; hr: high-risk; NILM: negative for intraepithelial lesions or malignancy; ASC-US: atypical squamous cells of undetermined significance; ASC-H: atypical squamous cells where HSIL cannot be excluded; LSIL: low-grade squamous intraepithelial lesions; HSIL: high-grade squamous intraepithelial lesions.

3.5. HPV vaccination effects

For the computation of the impact of HPV vaccination, weighted data including all enrolled women from both phases were used. Data on protection against infection with HPV16, HPV18, HPV6 or HPV11 and against cytological and histological lesions associated with HPV16 or 18 in women younger than 30 are shown in Table 2. Significant protection was observed regarding infection with HPV16 (VE = 67%, 95% CI: 48% to 79%), HPV18 (93%, 95% CI: 52–99%), HPV16/18 (71%, 95% CI: 55–81%) and HPV 11, VE = 100%, CI not computable, but CI around RD excluding zero. The protection against HPV6 infection was not significant (VE = 33%, 95% CI: −55% to 71%). Overall hrHPV infection was lower in vaccinated versus non-vaccinated women (VE = 16%, 95% CI: 2–29%). However, vaccinated women had a non-significantly higher risk of infection with other hrHPV without HPV16/18 than non-vaccinated women VE = −15% (95% CI: −38% to 4%). HPV vaccination protected also against cytological lesions associated with HPV16/18 and even against cytological lesions irrespective of presence of hrHPV (significant for ASC-US+ and LSIL+ associated with HPV16/18 and for overall LSIL+ and HSIL+). Finally, vaccination protected against histologically confirmed lesions, which resulted in significantly lower absolute risks of CIN1+ (RD = −1.6%, 95% CI: −2.6% to −0.7%) and CIN3+ associated with HPV16/18 (-0.3%, 95% CI: −0.6% to −0.1%).

Protection against HPV16 or HPV18 infection and against cytological lesions associated with these infections, separated by finer age groups are shown in Table 3. Vaccine effectiveness decreased with age. Protection was significant against HPV16 and 18 infection in all age groups, against ASC-US in age group 18–19 and against LSIL in age groups 18–19 and 20–24. A non-significant higher risk of HSIL associated with HPV16/18 was noted in vaccinated women aged 25–29 ($p = 0.3986$).

4. Discussion

The SEHIB study provides a solid basis of information on the current burden of HPV infection and related cervical precancer in Belgium and provides an estimate of the early impact of the first years of HPV vaccination among young women attending screening. Significant protection was observed regarding infection with the HPV types included in the vaccines as well as cytological and histological precancerous lesions associated with HPV16 and HPV18.

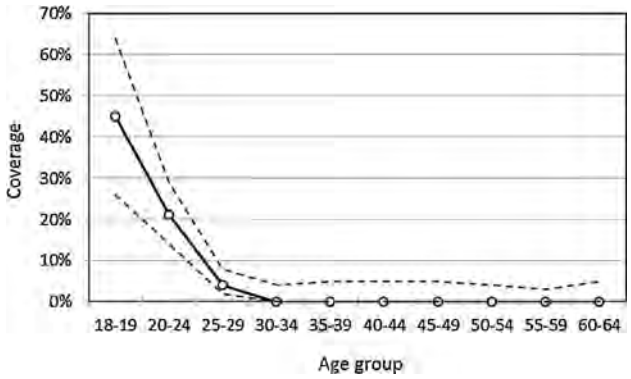


Fig. 3. HPV vaccination coverage (and 95% confidence interval [interrupted lines]), by age group, estimated in the screening population of the SEHIB study.

Table 2
Protection against infection with HPV16, HPV18, HPV6 or HPV11 and against cytological and histological lesions associated with HPV16 or 18, expressed as vaccine effectiveness and risk difference in women aged 18–29 years.

Protection against	Vaccine effectiveness	Risk difference
HPV16 infection	67% (95% CI: 48% to 79%)	-5.7% (95%CI: -7.2% to -4.1%)
HPV18 infection	93% (95% CI: 52% to 99%)	-1.9% (95%CI: -2.5% to -1.3%)
HPV1618 infection	71% (95% CI: 55% to 81%)	-7.0% (95%CI: -8.5% to -5.4%)
HPV6 infection	33% (95% CI: -55% to 71%)	-0.4% (95%CI: -1.2% to 0.4%)
HPV11 infection	100% (CI non computable)	-0.4% (95%CI: -0.6% to -0.2%)
hrHPV infection	16% (95% CI: 2% to 29%)	-3.8% (95%CI: -7.0% to -0.6%)
hrHPV other than 16/18 infection	-15% (95% CI: -38% to 4%)	2.2% (95%CI: -0.9% to 5.3%)
ASC-US+ associated with HPV16/18	60% (95% CI: 26% to 78%)	-2.2% (95%CI: -3.3% to -1.1%)
LSIL+ associated with HPV16/18	74% (95% CI: 34% to 90%)	-1.8% (95%CI: -2.6% to -1.0%)
HSIL+ associated with HPV16/18	41% (95% CI: -78% to 81%)	-0.3% (95%CI: -0.9% to 0.2%)
ASC-US irrespective of HPV types	11% (95% CI: -10% to 28%)	-1.5% (95%CI: -4.1% to 1.1%)
LSIL irrespective of HPV types	32% (95% CI: 6% to 50%)	-2.6% (95%CI: -4.5% to -0.7%)
HSIL irrespective of HPV types	63% (95% CI: 3% to 86%)	-1.0% (95%CI: -1.7% to -0.3%)
CIN1+ associated with HPV16/18	75% (95% CI: -4% to 94%)	-0.8% (95%CI: -1.4% to -0.3%)
CIN2+ associated with HPV16/18	54% (95% CI: -98% to 89%)	-0.3% (95%CI: -0.8% to 0.1%)
CIN3+ associated with HPV16/18	100% (95% CI non computable)	-0.2% (95%CI: -0.4% to -0.1%)
CIN1+ irrespective of HPV types	69% (95% CI: 14% to 89%)	-1.6% (95%CI: -2.6% to -0.7%)
CIN2+ irrespective of HPV types	63% (95% CI: -57% to 91%)	-0.6% (95%CI: -1.3% to 0.0%)
CIN3+ irrespective of HPV types	100% (95% CI non computable)	-0.3% (95%CI: -0.6% to -0.1%)

Dark green	significant protection
Light green	non-significant protection, VE>50%
Very light green	non-significant protection, VE<50%
Light yellow	non-significant higher risk
Yellow	significant higher risk

HPV: human papillomavirus; ASC-US+: atypical squamous cells of undetermined significance or worse, LSIL+: low-grade squamous intraepithelial lesion or worse, HSIL+: high-grade squamous intraepithelial lesion or worse; CIN1+: cervical intraepithelial neoplasia of grade 1 or worse.

4.1. Observed vaccination effects

The protection against infection with the vaccine types 16 and 18 observed in SEHIB (68 to 77% in the age groups 18–30, see Table 3) is comparable with the VE against HPV16/18 infection reported from randomised trials in the group of all included women regardless of HPV DNA status at enrolment (age group 15–26 years) [13]. Protection against incident HPV16/18 infection in this group is reported only for the bivalent vaccine: 76% (95% CI: 67–83%) [14]. The vaccine efficacy against 6-month persistent HPV16/18 infection was 56% (95% CI: 49–62%) [13], with two included trials assessing the bivalent vaccine where VE was 52% [15] and 59% [16], respectively. Although the SEHIB study could not distinguish between incident or persistent infection, we can conclude that the observed effectiveness against prevalent HPV16/18 is of the same order of magnitude as the protection against incident and persistent infection reported from the trials.

In the vaccine trials, the pooled VE against CIN2+ due to HPV16/18 in the total vaccinated group was 54% (95% CI: 43% to 63%) [13]. This pooled VE was estimated from three trials where VE was 48% [17] and 64% [18] for the quadrivalent and 60% [16] for the bivalent vaccine, respectively. The SEHIB VE estimate (54%) is within this range. The VE against CIN2+ irrespective of HPV types pooled from RCTs was 26% (95% CI: 13% to 37%) [13]. This pooled estimate was derived from three trials where VE was 19% [17], 33% [16] and 36% [18], respectively. The corresponding SEHIB VE estimate was again non-inferior (63%, but with large confidence intervals). In SEHIB, vaccine effectiveness was in general best in the 18–19 year age group and dropped with increasing age. In the 25–29 age group, HPV vaccination did not protect against cytological lesions due to HPV16/18. This finding is compatible with the higher probability that young adult women (who are beyond the target age for reimbursement) were vaccinated after exposure to HPV infection and that women with high-risk sexual behaviour may be more motivated to buy the vaccine than those with a lower risk profile [19]. It was already known that current prophylactic HPV vaccines do not clear existing infection with one of the vaccine types [20]. There was a slight non-significant increase in other non16/18-hrHPV infection among vaccinated women versus non vaccinated women (VE = −15%, 95% CI −38 to 4%). Although this phenomenon may be explained by chance ($p = 0.14$), type replacement, not detectable because of insufficient power, cannot be excluded. Nevertheless, no repercussion on protection against cytological or histological lesions irrespective of HPV types was noted in SEHIB. This latter finding may support the statement that type replacement – if it would occur – may not result in higher rates of cervical (pre) cancer because of the lower malignant potential of the other high-risk types. It is expected that the lower vaccine protection

Table 3
Protection against HPV16 or HPV18 infection and against cytological lesions associated with these infections, expressed as vaccine effectiveness and risk difference in age groups 18–19 years, 20–24 years, 25–29 years.

Protection against	Age group	Vaccine effectiveness	Risk difference
Infection with HPV16/18	18-19	77% (95% CI: 19% to 94%)	-10.2% (95%CI:-19.6 % to -0.7%)
	20-24	74% (95% CI: 52% to 85%)	-7.7% (95%CI:-10.3% to -5.1%)
	25-29	68% (95% CI: 25% to 86%)	-5.9% (95%CI: -8.6% to -3.3%)
ASCUS+ associated with HPV16/18	18-19	100% (CI not computable)	-7.4% (95%CI: -14.3% to -0.5%)
	20-24	75% (95% CI: -13% to 94%)	-4.4% (95%CI: -8.0% to -0.8%%)
	25-29	4% (95% CI: -127% to 59%)	-0.1% (95%CI: -2.6% to 2.3%)
LSIL+ associated with HPV16/18	18-19	100% (CI not computable)	-7.4% (95%CI: -14.3% to -0.5%)
	20-24	91% (34% to 99%)	-2.4% (95%CI: -3.5% to -1.2%)
	25-29	7% (95% CI: -175% to 68%)	-0.1% (95%CI: -2.1% to 1.8%)
HSIL+ associated with HPV16/18	18-19	100% (CI not computable)	-0.3% (95%CI: -1.0% to 0.4%)
	20-24	56% (95% CI: -287% to 95%)	-32.6% (95%CI:-20.7% to-27.9%)
	25-29	-76% (95% CI: -565% to 53%)	0.6% (95%CI: -1.1% to 2.2%)

Dark green	significant protection
Light green	non-significant protection, VE>50%
Very light green	non-significant protection, VE<50%
Light yellow	non-significant higher risk
Yellow	significant higher risk

HPV: human papillomavirus; ASC-US+: atypical squamous cells of undetermined significance or worse, LSIL+: low-grade squamous intraepithelial lesion or worse, HSIL+: high-grade squamous intraepithelial lesion or worse.

observed now in women older than 25 will disappear when more subjects vaccinated at school age will reach the screening group, since they belong to cohorts who received their vaccine doses before onset of sexual activity (exposure to HPV).

The SEHIB findings are in line with surveillance reports and trend analyses from other countries [21–25] and demonstrate the utility of HPV genotyping of archived cervical samples collected in the framework of routine cervical cancer screening to evaluate the impact of HPV vaccination [26].

4.2. Strengths and weaknesses of SEHIB

SEHIB confirms that the protective effects of HPV vaccines, demonstrated in randomised trials, also are reflected in protective effects observable in a real life screening setting. By using a weighting of records from continuous screening samples enriched with pathological samples, we were able to identify a number of cervical precancer lesions from a virtual screening population of ~20,000 women (increasing the statistical power) while genotyping only ~6000 samples (reducing testing costs to less than one third). The fact that the vaccination status was based on anamnestic data was one of the weaknesses of SEHIB. Nevertheless, validation by contacting the vaccinating physician in a first phase revealed a high level of concordance, allowing the interruption of this time consuming verification procedure. A recent survey among Australian young women also reported good agreement between the self-reported and administratively validated HPV vaccination status [27].

The suboptimal rate of biopsy retrieval has reduced the power for histological endpoints and even may have generated bias. Reasons for the low biopsy rate may be: biopsy in other gynaecological centres, lost-to-follow-up of patients and insufficient available tissue from small punch biopsies. In spite of the potential bias, we believe that its magnitude may be small given the consistency between cytological and histological outcomes.

4.3. Recommendations for further research and surveillance

Setting up comprehensive vaccination registries linked with screening and cancer registries and combinable with cervical cytology and histology biobanks would guaranty more reliable information [28]. Such comprehensive linked databases are crucial to verify the impact of HPV vaccination such as the evolution of the incidence of cervical and other HPV-associated cancers, effects of 2-dose vaccine schedules, the duration of protection as well as the occurrence of possible type-replacement and cross-protection [26,29]. Moreover, joined screening and vaccination registries may offer the framework to evaluate new experimental screening and vaccination combinations and will allow physicians to diversify future screening according to the vaccination history of the individual woman.

5. Conclusions

The SEHIB study constitutes a baseline of the current burden of HPV infection and associated cervical disease in Belgium and clearly demonstrates effectiveness of HPV vaccination in the community of women younger than 30 years. Reduction in HPV16 and HPV18 infection was strongest in the age group of 18–19 years and declined with increasing age.

Funding

M. Arbyn/JP Bogers received support from the 7th Framework Program of DG Research of the European Commission through the CoheaHr Project [Grant number 603019], coordinated by the VU University Medical Centre of Amsterdam, the Netherlands and the HPV-AHEAD Project, coordinated by IARC, Lyon, France [Grant number: FP7-HEALTH-2011-282562].

An unrestricted grant was provided by Sanofi-Pasteur-MSD to the University of Ghent who coordinated the SEHIB study. The

grant was given in the framework of the EMA (European Medicine Agency) request to set up post-marketing surveillance of HPV vaccination effects in non-Nordic member states of the European Union.

Acknowledgements

Edward Camerlinck, Bart Craeye, Aurore Guieu, Bavo Verpoest, Liesbeth Wildero-Van Wouwe and Dirk Van Braeckel are acknowledged for logistic and administrative support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.canep.2015.12.011.

References

[1] Belgian Cancer Registry Foundation. Belgisch Kankerregister/Registre du Cancer (incidence year 2011). http://www.kankerregister.org/. 2014.
[2] M. Arbyn, X. Castellsagué, S. de Sanjosé, et al., Worldwide burden of cervical cancer in 2008, Ann. Oncol. 22 (2011) 2675–2686.
[3] M. Arbyn, H. Van Oyen, Cervical cancer screening in Belgium, Eur. J. Cancer 36 (2000) 2191–2197.
[4] M. Arbyn, C. Simoens, H. Van Oyen, et al., Analysis of 13 million individual patient records pertaining to Pap smears, colposcopies, biopsies and surgery on the uterine cervix (Belgium, 1996–2000), Prev. Med. 48 (2009) 438–443.
[5] M. Arbyn, V. Fabri, M. Temmerman, C. Simoens, Attendance at cervical cancer screening and use of diagnostic and therapeutic procedures on the uterine cervix assessed from individual health insurance data (Belgium, 2002–2006), PLoS One 9 (2014) e92615.
[6] M. Arbyn, HPV vaccination of teenage girls in Belgium, Annual Conference; Bredene (Belgium), Flemish Society of Obstetric and Gynaecology, 2012, pp. 1–21.
[7] G. Top, Paeps A. Actualisering, HPV-vaccinatie in Vlaanderen—Resultaten van de eerste twee vaccinatiejaren 2010–2012, Vlaams Infectiebulletin 81 (2015) 1–2.
[8] A. Vermeeren, M.C. Miermans, B. Swennen, Évolution de 2008 à 2013 des couvertures vaccinales des enfants et jeunes en âge scolaire en Fédération Wallonie-Bruxelles, PROVAC, Brussels, 2014.
[9] M. Arbyn, A. Herbert, U. Schenck, et al., European guidelines for quality assurance in cervical cancer screening: recommendations for collecting samples for conventional and liquid-based cytology, Cytopathology 18 (2007) 133–139.
[10] D. Solomon, D. Davey, R. Kurman, et al., The 2001 Bethesda System: terminology for reporting results of cervical cytology, JAMA 287 (2002) 2114–2119.
[11] C.E. Depuydt, I.H. Benoy, J.F. Beert, A.M. Criel, J.J. Bogers, M. Arbyn, Clinical validation of a type-specific real time quantitative human papillomavirus PCR to the performance of Hybrid Capture 2 for the purpose of cervical cancer screening, J. Clin. Microbiol. 50 (2012) 4073–4077.
[12] I.M. Micalessi, G.A. Boulet, J.J. Bogers, I.H. Benoy, C.E. Depuydt, High-throughput detection, genotyping and quantification of the human papillomavirus using real-time PCR, Clin. Chem. Lab. Med. 50 (2012) 655–661.
[13] M. Arbyn, A. Bryant, P.L. Martin-Hirsch, L. Xu, C. Simoens, L. Markowitz, Prophylactic vaccination against human papillomaviruses to prevent cervical cancer and its precursors (Protocol), Cochrane Database Syst. Rev. 12 (2013) 1–15.
[14] A.R. Kreimer, P. Gonzalez, H.A. Katki, et al., Efficacy of a bivalent HPV 16/18 vaccine against anal HPV 16/18 infection among young women: a nested analysis within the Costa Rica vaccine trial, Lancet Oncol. 12 (2011) 862–870.
[15] R. Herrero, S. Wacholder, A.C. Rodriguez, et al., Prevention of persistent human papillomavirus infection by an HPV16/18 vaccine: a community-based randomized clinical trial in guanacaste, Costa Rica, Cancer Discovery 1 (2011) 408–419.
[16] M. Lehtinen, J. Paavonen, C.M. Wheeler, et al., Overall efficacy of HPV-16/18 AS04-adjuvanted vaccine against grade 3 or greater cervical intraepithelial neoplasia: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial, Lancet Oncol. 13 (2012) 89–99.
[17] N. Munoz, S.K. Kjaer, K. Sigurdsson, et al., Impact of human papillomavirus (HPV)-6/11/16/18 vaccine on all HPV-associated genital diseases in young women, J. Natl. Cancer Inst. 102 (2010) 325–339.
[18] A. Rowhani-Rahbar, C. Mao, J.P. Hughes, et al., Longer term efficacy of a prophylactic monovalent human papillomavirus type 16 vaccine, Vaccine 27 (2009) 5612–5619.
[19] A. Leval, E. Herweijer, A. Ploner, et al., Quadrivalent human papillomavirus vaccine effectiveness: a Swedish national cohort study, J. Natl. Cancer Inst. 105 (2013) 469–474.
[20] A. Hildesheim, R. Herrero, S. Wacholder, et al., Effect of human papillomavirus 16/18 L1 viruslike particle vaccine among young women with preexisting infection: a randomized trial, JAMA 298 (2007) 743–753.
[21] J.M. Brotherton, M. Fridman, C.L. May, G. Chappell, A.M. Saville, D.M. Gertig, Early effect of the HPV vaccination programme on cervical abnormalities in Victoria, Australia: an ecological study, Lancet 377 (2011) 2085–2092.
[22] S.N. Tabrizi, J.M. Brotherton, J.M. Kaldor, et al., Fall in human papillomavirus prevalence following a national vaccination program, J. Infect. Dis. 206 (2012) 1645–1651.
[23] B. Baldur-Felskov, C. Dehlendorff, C. Munk, S.K. Kjaer, Early impact of human papillomavirus vaccination on cervical neoplasia–nationwide follow-up of young Danish women, J. Natl. Cancer Inst. 106 (2014) djt460.
[24] K. Kavanagh, K.G. Pollock, A. Potts, et al., Introduction and sustained high coverage of the HPV bivalent vaccine leads to a reduction in prevalence of HPV 16/18 and closely related HPV types, Br. J. Cancer 110 (2014) 2804–2811.
[25] M. Merckx, D. Vanden Broeck, I. Benoy, C. Depuydt, M. Arbyn, Early effects of HPV vaccination in Belgium, Eur. J. Cancer Prev. 24 (2015) 340–342.
[26] J. Dillner, M. Arbyn, E. Unger, L. Dillner, Monitoring of human papillomavirus vaccination, Clin. Exp. Immunol. 163 (2011) 17–25.
[27] S.L. Osborne, S.N. Tabrizi, J.M. Brotherton, et al., Assessing genital human papillomavirus genoprevalence in young Australian women following the introduction of a national vaccination program, Vaccine 33 (2015) 201–208.
[28] M. Arbyn, K. Andersson, C. Bergeron, J.P. Bogers, M. von Knebel-Doeberitz, J. Dillner, Cervical cytology biobanks as a resource for molecular epidemiology, Methods Mol. Biol. 675 (2011) 279–298.
[29] M.A. Stanley, S.L. Sudenga, A.R. Giuliano, Alternative dosage schedules with HPV virus-like particle vaccines, Expert Rev. Vaccines 13 (2014) 1027–1038.

REVIEW

Annals of Internal Medicine

Genotyping for Human Papillomavirus Types 16 and 18 in Women With Minor Cervical Lesions

A Systematic Review and Meta-analysis

Marc Arbyn, MD, MSc, DrTMH; Lan Xu, MSc; Freija Verdoodt, PhD; Jack Cuzick, PhD; Anne Szarewski, MD, PhD†; Jerome L. Belinson, MD; Nicolas Wentzensen, MD, PhD; Julia C. Gage, PhD, MPH; and Michelle J. Khan, MD, MPH

Background: High-risk human papillomavirus (hrHPV) testing to triage women with minor cervical lesions generates many referrals.

Purpose: To evaluate the accuracy of genotyping for HPV types 16 and 18 and its utility as a second triage step after hrHPV testing in women with minor cervical lesions.

Data Sources: Searches of 4 bibliographic databases, without language restrictions, from 1 January 1999 to 1 February 2016.

Study Selection: Studies involving women with atypical squamous cells of undetermined significance (ASC-US) or low-grade squamous intraepithelial lesions (LSIL) who were triaged with tests for hrHPV and HPV 16/18 to find cervical intraepithelial neoplasia (grade ≥2 [CIN2+] or grade ≥3 [CIN3+]).

Data Extraction: Independent study selection, extraction of data, and quality assessment by 2 reviewers.

Data Synthesis: Twenty-four moderate- to good-quality studies involving 8587 women with ASC-US and 5284 with LSIL were found. The pooled sensitivity of HPV 16/18 genotyping for CIN3+ was about 70% for women with either ASC-US or LSIL. The pooled specificity (with a threshold of grade <2 CIN) was 83% (95% CI, 80% to 86%) for women with ASC-US and 76% (CI, 74% to 79%) for those with LSIL. The average risk for CIN3+ was 17% and 19% in HPV 16/18–positive women with ASC-US and LSIL, respectively, and was 5% in hrHPV-positive but HPV 16/18–negative women with either ASC-US or LSIL.

Limitation: Methodological and technical heterogeneity among studies; insufficient data to assess accuracy of separate assays.

Conclusion: Testing for HPV 16/18 to triage women with minor abnormal cytology is poorly sensitive but may be useful as a second triage test after hrHPV testing, with direct referral if the woman is positive for HPV 16/18. Whether colposcopy or repeated testing is recommended for hrHPV-positive but HPV 16/18–negative women depends on local decision thresholds that can be derived from pretest–posttest probability plots.

Primary Funding Source: 7th Framework Programme of the European Commission.

Ann Intern Med. 2017;166:118-127. doi:10.7326/M15-2735 **www.annals.org**

For author affiliations, see end of text.

This article was published at www.annals.org on 15 November 2016.

† Deceased.

Several countries have switched to human papillomavirus (HPV)–based screening for cervical cancer, but cytologic examination of a Papanicolaou (Pap) smear remains the primary form of cervical cancer screening in many other countries. Direct referral for diagnostic work-up with colposcopy and biopsy is usually recommended for women with high-grade lesions. However, women with minor cytologic abnormalities, including atypical squamous cells of undetermined significance (ASC-US) or low-grade squamous intraepithelial lesions (LSIL), have only a modestly increased risk for cervical cancer (1). In the past, repeating the Pap test (repeated cytology) was the recommended follow-up for women with ASC-US or LSIL. Given the strong etiologic link between high-risk HPV (hrHPV) infection and cervical cancer, hrHPV testing has been proposed as an alternative triage method for women with equivocal or mildly abnormal cytology.

Randomized trials and systematic reviews show that, compared with repeated cytology, hrHPV testing has higher sensitivity and similar specificity in identifying underlying or incipient cervical precancer in women with ASC-US (2–4). Accordingly, triage by hrHPV testing has become standard practice (5–8). Low-grade squamous intraepithelial lesions are associated with a risk for precancer similar to that among hrHPV-positive women with ASC-US (9). Because most women with LSIL test positive for hrHPV (10), triage by hrHPV testing is inefficient (11, 12). The widespread practice of referring all women with hrHPV infection and ASC-US or with LSIL to colposcopy carries a considerable burden and cost. Because HPV types 16 and 18 cause about 70% of cervical cancer cases (8), genotyping for these types has been proposed as an additional tool to allow more fine-tuned management.

In this article, we present the results of a systematic review on the accuracy of genotyping for HPV types 16 and 18 to triage women with ASC-US or LSIL, including those who are hrHPV-positive. We also present a framework to assess the clinical utility of triage tests, based on the risk for cervical precancer before and after triage.

See also:

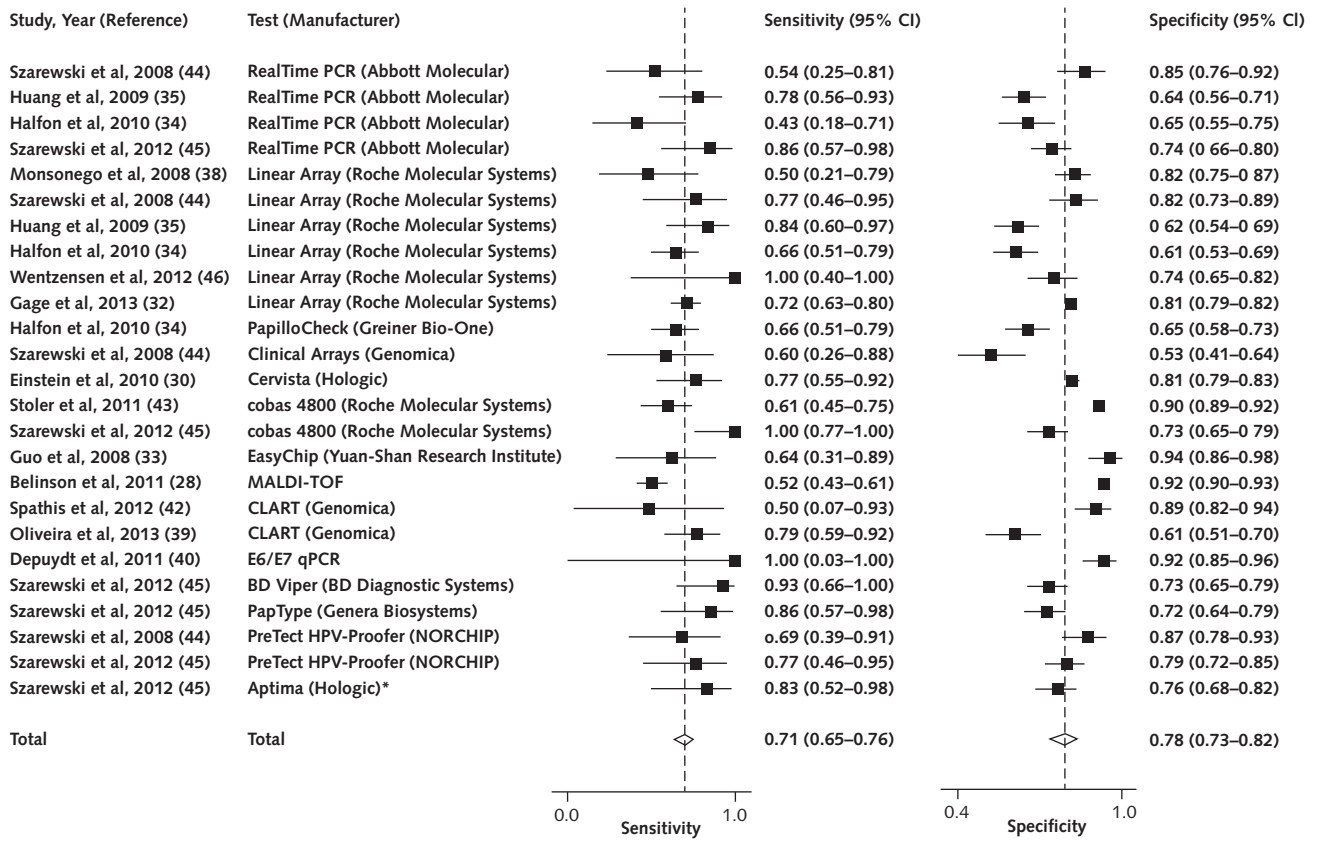
Web-Only
Supplement

Methods

We developed a protocol (**Supplement 1**, available at www.annals.org), followed standard procedures for

 © 2017 American College of Physicians

Figure 1. Meta-analysis of the sensitivity and specificity of genotyping for HPV 16/18 to detect CIN3+ in women with ASC-US.



ASC-US = atypical squamous cells of undetermined significance; CIN3+ = cervical intraepithelial neoplasia, grade ≥3; HPV = human papillomavirus; MALDI-TOF = matrix-assisted laser desorption/ionization time of flight; qPCR = quantitative polymerase chain reaction.
* Also included HPV 45 genotyping.

meta-analyses of diagnostic accuracy studies (12, 13), and reported processes and results according to standard guidelines (14).

Data Sources and Searches

Using the search strategies given in the protocol, we searched PubMed/MEDLINE, EMBASE, Scopus, and the Cochrane Central Register of Controlled Trials with no language restrictions from 1 January 1999 to 1 February 2016. We also culled reference lists of selected reports.

Study Selection

Two reviewers independently screened titles and abstracts to identify relevant studies. Studies had to involve 20 or more women with either ASC-US or LSIL who had cervical samples tested with an assay detecting hrHPV and HPV types 16 and 18, as well as a reference test to verify presence or absence of cervical intraepithelial neoplasia (grade ≥2 [CIN2+] or grade ≥3 [CIN3+]).

The ASC-US group comprised women with either ASC-US as defined according to the 1988 and 2001 editions of the Bethesda System (15, 16) or borderline dyskaryosis (17). The LSIL group included women with either LSIL (16) or mild dyskaryosis (17). Authors were contacted if no separate accuracy data were reported for ASC-US and LSIL or when CIN3+ was the only outcome reported.

Tests and Reference Standards

The evaluated index tests were assays identifying DNA or RNA of HPV types 16 and 18, jointly or separately (HPV 16/18). An HPV 16/18 test result was considered positive if HPV 16 or HPV 18 was present and negative if both types were absent. The comparator tests were hrHPV assays identifying at least 8 hrHPV types (HPV 16, 18, 31, 33, 35, 45, 52, and 58). The HC2 assay (Qiagen) was used as the hrHPV comparator test if present. In studies where the HC2 assay was not used, other hrHPV assays or genotyping tests identifying separate hrHPV types were accepted as the comparator test. Details on test platforms and the panel of considered hrHPV types were noted. The cutoff proposed by the manufacturer of each assay was used as the positivity criterion.

In addition to use of HPV 16/18 genotyping as the sole triage test, a combined triage strategy was assessed, where HPV 16/18 genotyping as a second step

Table 1. Pooled Absolute Sensitivity and Specificity of HPV 16/18 Genotyping in Triage of Women With ASC-US or LSIL to Detect Underlying CIN2+ or CIN3+

Outcome, by Triage Group	Study/Test Reports, *n*	Pooled Value (95% CI), %	
		Sensitivity	Specificity
Women with ASC-US			
CIN2+	32	58.8 (54.6–62.9)	82.9 (79.6–85.7)
CIN3+	25	70.7 (64.9–76.0)	78.1 (73.3–82.3)
Women with LSIL			
CIN2+	28	55.5 (52.4–58.5)	76.3 (73.5–78.9)
CIN3+	24	70.0 (65.4–74.2)	72.5 (69.0–75.8)

ASC-US = atypical squamous cells of undetermined significance; CIN2+ = cervical intraepithelial neoplasia, grade ≥2; CIN3+ = cervical intraepithelial neoplasia, grade ≥3; HPV = human papillomavirus; LSIL = low-grade squamous intraepithelial lesions.

was restricted to women who were hrHPV-positive at the first triage test. All women underwent verification with colposcopy, colposcopy-directed biopsies (possibly supplemented with random biopsies), or endocervical curettage. The type of verification (reference standard) was recorded for each study. Two levels of the disease outcome were considered: CIN2+ and CIN3+. Adenocarcinoma in situ was included in the CIN3+ outcome.

Data Extraction and Quality Assessment

Two authors (M.A. and M.J.K.) independently checked the eligibility of references and extracted the number of true-positive, true-negative, false-positive, and false-negative results for each test, triage group, and outcome. Information on the study design, the clinical setting where patients were enrolled, the HPV assays, and the verification procedures was condensed into comprehensive tables. The quality of the selected studies was evaluated independently by 2 coauthors (L.X., F.V., or M.J.K.) using the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) checklist (18, 19).

Data Synthesis and Analysis

The absolute sensitivity and specificity were pooled by using a bivariate normal model for the logit transformations of sensitivity and specificity (20, 21). Summary receiver-operating characteristic curves and forest plots were created to show the joint overall and study-specific sensitivity and specificity of HPV 16/18 genotyping for triage of women with ASC-US or LSIL.

The relative sensitivity and specificity of the index tests versus the comparator tests were computed by including the test as a covariate in the bivariate model (22, 23). We used the Linear Array assay (Roche Molecular Systems) as the comparator to assess the relative accuracy of HPV 16/18 genotyping with different assays. Sources of heterogeneity of accuracy were assessed by including a series of potentially influential covariates in the bivariate model: QUADAS items, type of gold standard, HPV test platform used for hrHPV testing or HPV 16/18 genotyping, target of the HPV test (DNA or RNA), and number of HPV types targeted by the hrHPV test. The Deeks regression test, based on the regression of the log diagnostic odds ratio onto 1/(effective sample size), was used to assess small-study effects (publication bias) (24). Statistical significance was defined as a *P* value less than 0.05. However, for the assessment of the variation of accuracy over multiple categories, we applied a Bonferroni correction (0.05/k, where k was the number of categories) to adjust the significance level. We conducted the statistical analyses with Stata, version 13 (StataCorp), and SAS Enterprise Guide, version 5.1 (SAS Institute).

Pretest–posttest probability plots were constructed to help evaluate the utility of tests and testing strategies. Posttest risks were computed from the average prevalence of precancer in the reviewed studies. Decision thresholds were based on benchmark risk levels applied in Europe (1% and 10%) and the United States (2.6% and 5.2%) (25–27).

Role of the Funding Source

This review was funded by the Directorate-General for Research & Innovation of the European Commission (7th Framework Programme, grant no. 603019). The funders had no role in designing, conducting, or reporting the review.

RESULTS

Selection of Studies and Study Characteristics

From 899 references, 24 studies were selected that met inclusion criteria (Figure 1 of Supplement 2, available at www.annals.org). The studies involved 8587 women with ASC-US and 5284 with LSIL (28–51). An overview of study designs, populations, and test characteristics is provided in Tables 1 and 2 of Supplement 2. Additional data were obtained from the authors of most of the studies, except for 8 articles that contained all of the absolute numbers required for computation of sensitivity and specificity (29, 30, 37, 47–51). Study settings included colposcopy clinics (30, 32, 34–36, 38, 40–42, 44–46, 49–51), primary screening settings (28, 29, 39, 43, 47), a maternity center (31), and pathology archives (33, 37). Fifteen HPV assays were evaluated.

Quality of Included Studies

Table 3 of Supplement 2 summarizes the methodological quality of the included studies. Most studies were of moderate or good quality; 3 were scored as probably free of bias (29, 32, 43), and 1 was scored as poor-quality (37). The most common potential sources of bias were failure to report on uninterpretable or equivocal test results (*n* = 12), failure to account for withdrawals (*n* = 9), and an unclear delay between tests (*n* = 8).

Absolute Sensitivity and Specificity

The summary receiver-operating characteristic curves (Figure 2 of Supplement 2) and forest plots (Figure 1) show the variation among studies and the pooled values of the sensitivity and specificity of genotyping for HPV 16/18 to detect CIN2+ or CIN3+ in women with ASC-US and LSIL. Genotyping for HPV

Table 2. Variation in Sensitivity and Specificity of HPV 16/18 Genotyping to Detect CIN2+ in Women With ASC-US or LSIL, by Test System

Test (Manufacturer)	Studies, *n*	Women With ASC-US				Women With LSIL			
		Sensitivity Ratio (95% CI)	*P* Value	Specificity Ratio (95% CI)	*P* Value	Sensitivity Ratio (95% CI)	*P* Value	Specificity Ratio (95% CI)	*P* Value
Linear Array (Roche Molecular Systems)*	10	1.00		1.00		1.00		1.00	
RealTime PCR (Abbott Molecular)	5	0.85 (0.61–1.17)	0.31	0.99 (0.93–1.05)	0.73	0.88 (0.64–1.19)	0.40	1.02 (0.97–1.07)	0.51
PapilloCheck (Greiner Bio-One)	1	0.97 (0.59–1.56)	0.87	1.02 (0.85–1.22)	0.85	1.03 (0.61–1.73)	0.92	1.08 (0.92–1.27)	0.35
Clinical Arrays (Genomica)	1	0.96 (0.52–1.79)	0.90	0.63 (0.50–0.81)	<0.001†	1.01 (0.69–1.48)	0.96	0.82 (0.71–0.95)	0.008
Cervista (Hologic)	2	1.14 (0.78–1.65)	0.50	1.00 (0.90–1.12)	0.95	–		–	
cobas 4800 (Roche Molecular Systems)	3	1.13 (0.92–1.39)	0.23	1.03 (0.97–1.08)	0.34	0.92 (0.74–1.14)	0.46	1.01 (0.96–1.07)	0.65
EasyChip (Yuan-Shan Research Institute)	1	1.09 (0.64–1.86)	0.74	1.16 (1.08–1.25)	<0.001†	0.78 (0.40–1.55)	0.48	1.17 (1.02–1.36)	0.030
MALDI-TOF	1	0.77 (0.46–1.30)	0.33	1.12 (1.00–1.21)	0.004	0.88 (0.64–1.21)	0.43	1.19 (1.09–1.29)	<0.001†
CLART (Genomica)	2	0.99 (0.64–1.54)	0.96	0.99 (0.89–1.12)	0.99	1.17 (0.89–1.54)	0.27	0.87 (0.76–1.01)	0.061
E6/E7 qPCR	1	1.39 (0.72–2.69)	0.33	1.12 (1.00–1.22)	0.004	0.60 (0.17–2.17)	0.44	0.95 (0.77–1.18)	0.66
BD Viper (BD Diagnostic Systems)	1	0.98 (0.15–6.32)	0.98	0.97 (0.73–1.29)	0.81	0.91 (0.59–1.42)	0.69	1.03 (0.91–1.16)	0.65
PapType (Genera Biosystems)	1	0.89 (0.13–6.30)	0.91	0.96 (0.72–1.29)	0.80	0.94 (0.61–1.45)	0.77	1.03 (0.91–1.17)	0.65
PreTect HPV-Proofer (NORCHIP)	2	0.96 (0.63–1.47)	0.85	1.05 (0.90–1.13)	0.27	0.86 (0.61–1.21)	0.39	1.10 (1.03–1.18)	0.003†
Aptima (Hologic)	1	0.85 (0.11–6.60)	0.88	1.00 (0.77–1.31)	0.98	0.91 (0.57–1.45)	0.68	1.00 (0.87–1.14)	0.94

ASC-US = atypical squamous cells of undetermined significance; CIN2+ = cervical intraepithelial neoplasia, grade ≥2; HPV = human papillomavirus; LSIL = low-grade squamous intraepithelial lesions; MALDI-TOF = matrix-assisted laser desorption/ionization time of flight; qPCR = quantitative polymerase chain reaction.
* Comparator test.
† Significant likelihood ratio test, which assesses whether the relative accuracy differs statistically from unity, with a significance level defined at 0.05/k (k = 14 [number of compared assays]; Bonferroni correction for multiple comparisons).

16/18 identified, on average, 70.7% (95% CI, 64.9% to 76.0%) of CIN3+ cases in women with ASC-US and 70.0% (CI, 65.4% to 74.2%) in women with LSIL (Table 1). The sensitivity of HPV 16/18 genotyping for CIN2+ was lower (difference of 12% to 14%) than for CIN3+. The pooled specificity to exclude CIN2+ was 82.9% (CI, 79.6% to 85.7%) in women with ASC-US and 76.3% (CI, 73.5% to 78.9%) in those with LSIL.

Table 2 shows the variation in the accuracy of genotyping for HPV 16/18 by test system, with the Linear Array assay as the comparator. No statistically significant differences in sensitivity were observed. However, a higher specificity ($P < 0.001$) was noted for EasyChip (Yuan-Shan Research Institute) in triage of women with ASC-US and for PreTect HPV-Proofer (NORCHIP) and MALDI-TOF (matrix-assisted laser desorption/ionization time of flight) in triage of women with LSIL, whereas a lower specificity was noted for Clinical Arrays (Genomica) in triage of women with ASC-US.

Relative Accuracy of Genotyping for HPV 16/18 Versus hrHPV Testing

The relative sensitivity of HPV 16/18 genotyping compared with hrHPV testing for detecting CIN3+ was 0.75 (CI, 0.68 to 0.83) in women with ASC-US and 0.70 (CI, 0.63 to 0.77) in those with LSIL (Table 3). The specificity of HPV 16/18 genotyping to exclude CIN2+ was substantially higher than for hrHPV testing (1.70 [CI, 1.51 to 1.90] in women with ASC-US and 3.14 [CI, 2.83 to 3.48] in those with LSIL). Results for HPV 16 genotyping compared with hrHPV testing and compared with HPV 16/18 genotyping are provided in Table 5 and Figure 5 of Supplement 2.

Table 3. Meta-analysis of the Relative Sensitivity and Specificity of HPV 16/18 Genotyping Versus High-Risk HPV Testing

Outcome, by Triage Group	Comparisons, *n*	Relative Sensitivity (95% CI)	*P* Value	Relative Specificity (95% CI)	*P* Value
Women with ASC-US					
CIN2+	29	0.59 (0.54–0.65)	<0.001	1.70 (1.51–1.90)	<0.001
CIN3+	15	0.75 (0.68–0.83)	<0.001	1.87 (1.64–2.12)	<0.001
Women with LSIL					
CIN2+	19	0.56 (0.51–0.62)	<0.001	3.14 (2.83–3.48)	<0.001
CIN3+	15	0.70 (0.63–0.77)	<0.001	3.49 (3.01–4.05)	<0.001

ASC-US = atypical squamous cells of undetermined significance; CIN2+ = cervical intraepithelial neoplasia, grade ≥2; CIN3+ = cervical intraepithelial neoplasia, grade ≥3; HPV = human papillomavirus; LSIL = low-grade squamous intraepithelial lesions.

Table 4. Accuracy and Pretest and Posttest Probabilities of CIN2+ and CIN3+ for Triage With hrHPV Testing or HPV 16/18 Genotyping

Outcome, by Triage Group	Pretest Risk, %*	Pooled Sensitivity, %	Pooled Specificity, %	Likelihood Ratio		Test Positivity Rate, %	Posttest Risk, %	
				Positive	Negative		If Result Is Positive	If Result Is Negative†
Women with ASC-US								
hrHPV testing								
CIN2+	10.1	95.0	48.6	1.85	0.10	55.8	17.2	1.1
CIN3+	6.0	96.6	45.0	1.76	0.08	57.5	10.1	0.5
HPV 16/18 genotyping								
CIN2+	10.1	58.8	82.9	3.43	0.50	21.3	27.7	5.3
CIN3+	6.0	70.7	78.1	3.23	0.38	24.8	16.9	2.4
hrHPV-positive women with ASC-US								
HPV 16/18 genotyping								
CIN2+	17.2	60.1	67.3	1.84	0.59	37.4	27.5	11.0
CIN3+	10.1	73.8	61.7	1.93	0.42	41.9	17.9	4.5
Women with LSIL								
hrHPV testing								
CIN2+	21.1	96.9	24.8	1.29	0.12	79.7	25.6	3.4
CIN3+	8.6	97.7	22.4	1.26	0.10	79.3	10.6	1.0
HPV 16/18 genotyping								
CIN2+	21.1	55.5	76.3	2.34	0.58	30.4	38.5	13.5
CIN3+	8.6	70.0	72.5	2.55	0.41	31.1	19.3	3.8
hrHPV-positive women with LSIL								
HPV 16/18 genotyping								
CIN2+	25.6	60.2	64.6	1.70	0.62	41.7	36.9	17.5
CIN3+	10.6	72.6	60.7	1.85	0.45	42.8	18.0	5.1

ASC-US = atypical squamous cells of undetermined significance; CIN2+ = cervical intraepithelial neoplasia, grade ≥2; CIN3+ = cervical intraepithelial neoplasia, grade ≥3; HPV = human papillomavirus; hrHPV = high-risk human papillomavirus; LSIL = low-grade squamous intraepithelial lesions.
* Average of the prevalence or short-term cumulative incidence of CIN2+ or CIN3+ pooled from the studies included in the meta-analysis. For 2-step triage, the pretest risk corresponds to the posttest risk after hrHPV testing.
† 1 – negative predictive value.

Influence of Study and Test Characteristics and Small-Study Effects

Few study-quality items influenced the accuracy of HPV 16/18 genotyping to detect underlying precancer (Table 6 of Supplement 2). In women with ASC-US, the specificity of HPV 16/18 genotyping was higher when withdrawal of cases was unclear and in cases of partial verification. In women with LSIL, the sensitivity was higher and the specificity was lower when an inappropriate reference test was used or when withdrawal of cases was not explained.

Genotyping for HPV 16/18 in triage of women with ASC-US was less sensitive when the reference standard involved additional random biopsies or a mix of gold standard tests, and was more specific when only 1 biopsy specimen from the most suspect area was taken compared with when multiple colposcopy-targeted biopsy specimens were taken (Table 7 of Supplement 2). In general, the choice of hrHPV testing platform (HC2 or other platform, panel of targeted hrHPV types, or DNA or RNA testing) did not influence the relative accuracy of HPV 16/18 genotyping compared with hrHPV testing. However, in triage of women with LSIL, the magnitude of the relative specificity was higher with the HC2 assay than with other hrHPV test platforms as comparators (Table 8 of Supplement 2). The relative accuracy of an RNA-based assay targeting HPV types 16/18/45 versus hrHPV testing did not differ from that of an assay targeting HPV 16/18 (Table 9 of Supplement 2).

The Deeks regression test for funnel plot asymmetry did not reveal small-study effects (Table 10 and Figure 6 of Supplement 2).

Meta-analysis of the Accuracy of HPV 16/18 Genotyping in hrHPV-Positive Women With ASC-US or LSIL

For women with ASC-US or LSIL who tested positive for hrHPV in a first-step triage test, the pooled sensitivity was similar or slightly higher and the specificity was lower (Table 11 and Figure 7 of Supplement 2) compared with all women with ASC-US or LSIL (Table 1).

Pretest and Posttest Risks for Cervical Precancer

Table 4 displays accuracy measures and pretest and posttest probabilities of CIN2+ and CIN3+. Before triage testing, the average pretest risk for CIN3+ was 6% among women with ASC-US. A positive hrHPV test result increased the average risk to 10.1%, and a negative result decreased the risk to 0.5%. A positive HPV 16/18 test result in women with ASC-US increased the risk to 16.9%, and a negative result decreased it to

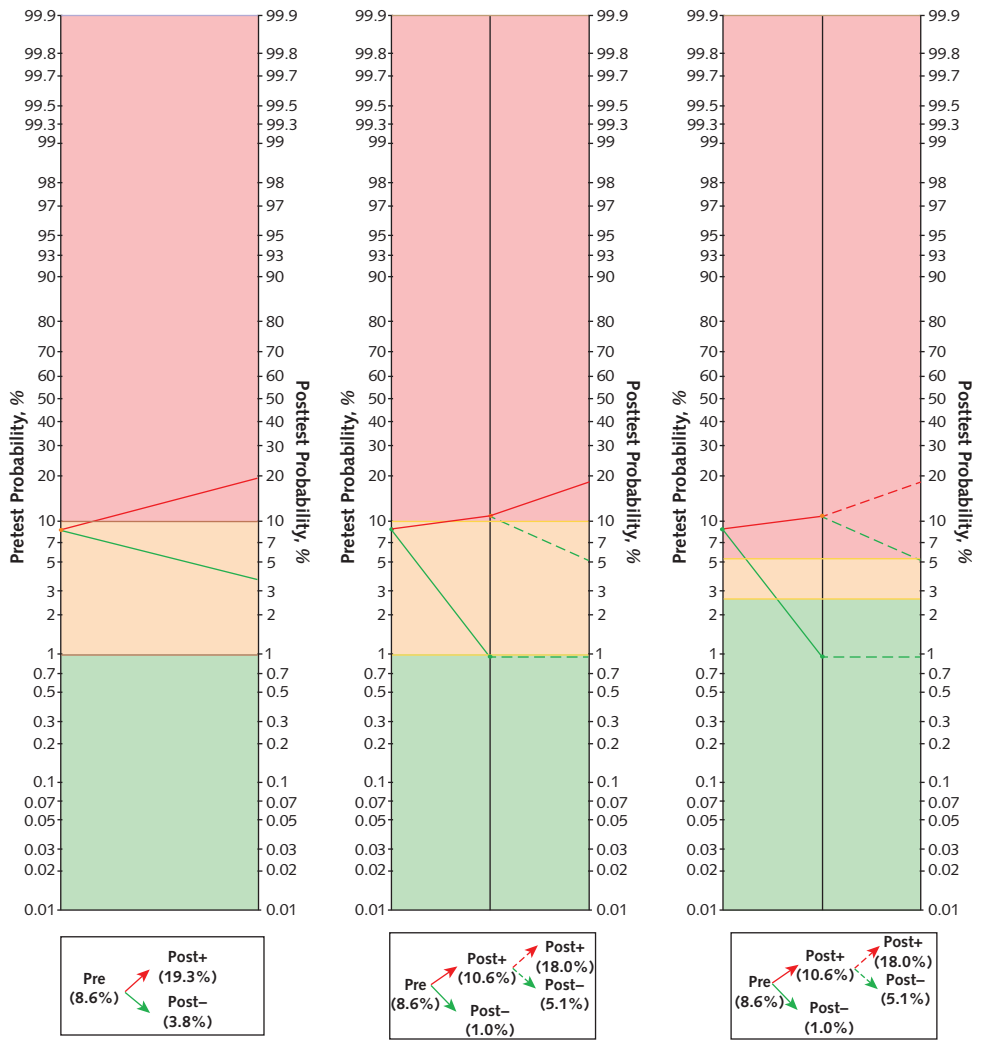
2.4%. In women with LSIL, the pretest probability of CIN3+ was 8.6%, whereas the posttest probabilities after triage were 10.6% in hrHPV-positive women, 19.3% in HPV 16/18–positive women, 1.0% in hrHPV-negative women, and 3.8% in HPV 16/18–negative women. With the 2-step triage, the risks for CIN3+ among women with ASC-US were 17.9% for those who were HPV 16/18–positive and 4.5% for those who were HPV 16/18–negative; for women with LSIL, the corresponding risks were 18.0% and 5.1%.

Pretest–Posttest Probability Plots

The utility of scenarios to triage women with LSIL, using the European (1% and 10%) and U.S. (2.6% and 5.2%) decision thresholds, is displayed in **Figure 2**. A positive HPV 16/18 test result would reclassify 30% of patients with LSIL as needing a colposcopy referral (**Figure 2**, *left* [*red zone*]), whereas the other 70% would need repeated testing (**Figure 2**, *left* [*yellow zone*]). The middle plot in **Figure 2** displays shifts in risk when hrHPV testing is followed by HPV 16/18 genotyping. Testing for hrHPV minimally increases the risk for CIN3+ (from 9% to 11%) for hrHPV-positive women with LSIL but decreases the risk to 1% for hrHPV-negative women with LSIL. Adding HPV 16/18 genotyping to hrHPV-positive women with LSIL would reclassify 43% as needing immediate referral (**Figure 2**, *middle* [*red zone*]). Repeated surveillance testing would be recommended for the remaining 57% who carry other hrHPV types (**Figure 2**, *middle* [*yellow zone*]).

In the United States, the pretest risk for CIN3+ in women with LSIL is considered sufficiently high to justify referral (**Figure 2**, *right*). Women who are negative for hrHPV could be followed conservatively (*green zone*), whereas those who are positive for HPV 16/18

Figure 2. Pretest and posttest probabilities of CIN3+ after triage in women with LSIL.



Use of HPV 16/18 genotyping as a single triage test (*left*) and use of a 2-step triage with hrHPV testing followed by HPV 16/18 genotyping in hrHPV-positive women (*middle* and *right*). Benchmarks are defined at risk levels of 1% and 10% (applied in Europe) (*left* and *middle*) and 2.6% and 5.2% (applied in the United States) (*right*). CIN3+ = cervical intraepithelial neoplasia, grade ≥3; HPV = human papillomavirus; hrHPV = high-risk human papillomavirus; LSIL = low-grade squamous intraepithelial lesions.

would require referral (**Figure 2**, *red zone*). Women who are hrHPV-positive but HPV 16/18–negative could be recommended for either colposcopy or retesting because their risk level is borderline (**Figure 2**, *between red and yellow zones*). Additional plots for diverse triage situations are shown in **Figures 8** to **13** of **Supplement 2**.

DISCUSSION

This meta-analysis found that genotyping for HPV 16/18 detects about 7 of 10 CIN3+ cases and about 6 of 10 CIN2+ cases in women with minor abnormal cytology. The pooled specificity (with a threshold of grade <2 CIN) was 83% in women with ASC-US and 76% in those with LSIL. Genotyping for HPV 16/18 was substantially more specific but less sensitive than testing for hrHPV. The average risks for underlying CIN3+ were 17% and 19% in HPV 16/18–positive women with ASC-US or LSIL, respectively; 2% and 4% among HPV 16/18–negative women with ASC-US or LSIL, respectively; and 5% in hrHPV-positive but HPV 16/18–negative women with either ASC-US or LSIL.

These findings, particularly the posttest risks for precancer, are useful for deciding how to incorporate HPV 16/18 genotyping results into patient management. Until recently, genotyping was done after hrHPV testing on women who tested positive. Several platforms are now available that allow inexpensive, high-throughput, 1-step genotyping. These platforms often give a readout of HPV 16/18 genotyping results separate from the other hrHPV testing, allowing the clinician immediate access to a secondary triage test. Our findings suggest that such partial genotyping tests can be used for risk stratification of hrHPV-positive women to immediate colposcopy or delayed follow-up. Local decision thresholds should help inform decisions about the clinical usefulness of the secondary triage strategy. According to European guidelines, a risk for CIN3+ greater than 10% is considered the threshold for referring a woman to colposcopy. In the United States, this threshold is greater than 5.2%; surveillance testing 6 to 12 months later is proposed if the risk for CIN3+ is 2.6% to 5.2%, and testing at longer intervals is proposed if the risk is less than 2.6% (25, 26). Women with ASC-US or LSIL who are hrHPV-positive but HPV 16/18–negative have a risk for underlying CIN3+ of around 5%, which U.S. guidelines would classify as borderline. This means that both immediate referral to colposcopy and retesting would be plausible options, with no clear preference (**Figure 2**, *right*). The utility of genotyping is more obvious in a European setting, where delayed retesting could be proposed for women with minor cytologic abnormalities who carry high-risk types other than HPV 16 and 18.

A strength of this meta-analysis is the inclusion of more than 8000 women with ASC-US and more than 5000 with LSIL from 24 studies. Our group previously reviewed the utility of hrHPV testing in triage of women with borderline and low-grade cytologic abnormalities (3, 4, 12, 52), but no group has performed a systematic review of HPV 16/18 genotyping as a primary or secondary triage test. This study helps clinicians to understand the underlying risks associated with HPV 16/18 positivity, which is now routinely reported in many newer HPV testing platforms.

In addition to our meta-analysis, we assessed the accuracy of triage using genotyping only for HPV 16, the most carcinogenic type. Genotyping for HPV 16/18 was 8% more sensitive for CIN3+ in both women with ASC-US and those with LSIL but 5% and 8% less specific, respectively, compared with HPV 16 genotyping (**Figure 5** and **Table 5** of **Supplement 2**).

Only 8 of the 24 included studies contained all of the required data in the published reports, but we obtained the data from the other studies directly from the authors. The studies were of moderate to good methodological quality, giving us confidence in the reliability of our sensitivity and specificity estimates. We found no evidence of publication bias or small-study effects. We found consistent and precise estimates of all relative accuracy estimates of HPV 16/18 genotyping compared with hrHPV testing.

Our review identified several limitations in the available data, including the lack of age-stratified data and too few data to precisely assess the accuracy of separate assays. Studies of test accuracy are observational in design, with short-term outcomes that do not provide evidence on effectiveness with respect to prevention of cancer (53). Randomized trials assessing cumulative incidence of CIN3+ or cancer among women who are negative at triage are needed to provide high-quality evidence on the efficacy of alternative management options. Studies had methodological and technical heterogeneity, although study quality and test characteristics seemed to have limited influence on estimates of test accuracy.

In conclusion, triage of women who have minor cytologic abnormalities with partial genotyping for HPV 16/18 increases efficiency compared with hrHPV testing, but at the expense of a loss in sensitivity. Whether a test has good triaging capability can be shown by plotting risks for precancer on pretest and posttest probability plots. Women testing positive for HPV 16 and 18 are at high risk and should be referred for colposcopy. Women carrying other high-risk HPV types cannot be released to routine screening. Whether the risk is sufficiently low in these women to avoid referral to colposcopy or to propose repeated testing depends on local decision thresholds.

From Scientific Institute of Public Health, Brussels, Belgium; Queen Mary University of London, London, United Kingdom; Cleveland Clinic, Cleveland, Ohio; National Cancer Institute, Bethesda, Maryland; and Kaiser Permanente Northern California, San Leandro, California.

Acknowledgment: The authors thank the following researchers for contributing requested unpublished data: Maria Benevolo (Pathology Department, Regina Elena National Cancer Institute, Rome, Italy) (41), Sonia Andersson (Department of Women's and Children's Health, Division of Obstetrics and

Gynecology, Karolinska Institutet, Stockholm, Sweden) (31), Ming Guo (Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, Texas) (33), Philippe Halfon (Virological Laboratory and Infectious Diseases, Laboratory Alphabio, Marseille, France) (34), Lee-Wen Huang (Department of Obstetrics and Gynecology, Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan) (54), Shihai Huang (Abbott Molecular, Des Plaines, Illinois) (35), François Coutlée (Département de Microbiologie et Infectiologie and de Gynécologie-Obstétrique et Pathologie, Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada) (36), Jean Monsonego (Institute of the Cervix, Paris, France) (38), Angela Pista (National Institute of Health, Lisbon, Portugal) (39), Aris Spathis (Department of Cytopathology, University General Hospital ATTIKON, School of Medicine, National and Kapodistrian University of Athens, Chaidari, Greece) (42), and Marc H. Stoler (University of Virginia Health System, Charlottesville, Virginia) (43). Additional unpublished data were provided by the following coauthors: Jerome Belinson (Preventive Oncology International, Cleveland, Ohio) (28), Julia Gage (Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, Maryland) (32), Jack Cuzick (Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, United Kingdom) (44), and Nicolas Wentzensen (Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland) (46).

Financial Support: Drs. Arbyn and Verdoodt and Ms. Xu were supported by the Belgian Foundation Against Cancer, Brussels, Belgium; the European Commission through the COHEAHR Network (grant no. 603019), coordinated by the Free University of Amsterdam (Amsterdam, the Netherlands), funded by the 7th Framework Programme of DG Research (Brussels, Belgium); the Joint Action CANCON, which has received funding from the European Union in the framework of the Health Programme (2008-13); and the Medizinische Hochschule of Hannover (Hannover, Germany). Dr. Khan was supported by training grant 5T32AI065388-05 from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health, institutional grant UL1 RR024131 through the National Institutes of Health/National Center for Research Resources University of California San Francisco–Clinical and Translational Science Institute Strategic Opportunities Support program, and mentored career development award K12 HS023009 through the Agency for Healthcare Research and Quality.

Disclosures: Dr. Cuzick reports grants from Qiagen, BD, Abbott, Hologic, Trovagene, Genera, and Cepheid; personal fees from Merck; and nonfinancial support from BD, Abbott, Hologic, Trovagene, and Cepheid during the conduct of the study. Dr. Belinson reports grants from Hologic and nonfinancial support from BGI Shenzhen during the conduct of the study. Dr. Wentzensen reports that he is an employee of the National Cancer Institute, which has received cervical cancer screening assays in-kind or at reduced cost from BD, Cepheid, Hologic, and Roche. Dr. Gage reports receiving human papillomavirus tests for research at no cost from Roche and BD. Dr. Khan reports travel reimbursement from Cepheid and the American Society for Colposcopy and Cervical Pathology outside the submitted work. Authors not named here have disclosed no conflicts of interest. Disclosures can also be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M15-2735.

Reproducible Research Statement: *Study protocol:* Available in **Supplement 1**. *Statistical code:* See the Methods and the study protocol. *Data set:* Available from Dr. Arbyn (e-mail, marc.arbyn@wiv-isp.be).

Requests for Single Reprints: Marc Arbyn, MD, MSc, DrTMH, Coordinator of the Unit of Cancer Epidemiology, Belgian Cancer Centre, Scientific Institute of Public Health, J. Wytsmanstreet 14, B1050 Brussels, Belgium; e-mail, marc.arbyn@wiv-isp.be.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Holowaty P, Miller AB, Rohan T, To T.** Natural history of dysplasia of the uterine cervix. J Natl Cancer Inst. 1999;91:252-8. [PMID: 10037103]
2. **ASCUS-LSIL Triage Study (ALTS) Group.** Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. Am J Obstet Gynecol. 2003;188:1383-92. [PMID: 12824967]
3. **Arbyn M, Buntinx F, Van Ranst M, Paraskevaidis E, Martin-Hirsch P, Dillner J.** Virologic versus cytologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. J Natl Cancer Inst. 2004;96:280-93. [PMID: 14970277]
4. **Arbyn M, Ronco G, Anttila A, Meijer CJ, Poljak M, Ogilvie G, et al.** Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. Vaccine. 2012;30 Suppl 5:F88-99. [PMID: 23199969] doi:10.1016/j.vaccine.2012.06.095
5. **Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al; ACS-ASCCP-ASCP Cervical Cancer Guideline Committee.** American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin. 2012;62:147-72. [PMID: 22422631] doi:10.3322/caac.21139
6. **Provencher DM, Murphy KJ.** Canadian consensus guidelines on human papillomavirus: chapter 3: the role of HPV testing. J Obstet Gynaecol Can. 2008;8:15-23.
7. **Medical Services Advisory Committee.** Human Papillomavirus Triage Test for Women with Possible or Definite Low-Grade Squamous Intraepithelial Lesions. Vol. 39. Canberra: Medical Services Advisory Committee; 2009.
8. **Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, et al.** European guidelines for quality assurance in cervical cancer screening. Second edition–summary document. Ann Oncol. 2010; 21:448-58. [PMID: 20176693] doi:10.1093/annonc/mdp471
9. **ASCUS-LSIL Triage Study (ALTS) Group.** A randomized trial on the management of low-grade squamous intraepithelial lesion cytology interpretations. Am J Obstet Gynecol. 2003;188:1393-400. [PMID: 12824968]
10. **Zuna RE, Wang SS, Rosenthal DL, Jeronimo J, Schiffman M, Solomon D; ALTS Group.** Determinants of human papillomavirus-negative, low-grade squamous intraepithelial lesions in the Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesions Triage Study (ALTS). Cancer. 2005;105: 253-62. [PMID: 15991244]
11. **Arbyn M, Martin-Hirsch P, Buntinx F, Van Ranst M, Paraskevaidis E, Dillner J.** Triage of women with equivocal or low-grade cervical cytology results: a meta-analysis of the HPV test positivity rate. J Cell MolMed.2009;13:648-59.[PMID:19166485]doi:10.1111/j.1582-4934.2008.00631.x

12. **Arbyn M, Roelens J, Simoens C, Buntinx F, Paraskevaidis E, Martin-Hirsch PP, et al.** Human papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions. Cochrane Database Syst Rev. 2013:CD008054. [PMID: 23543559] doi:10.1002/14651858.CD008054.pub2
13. **Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P.** Cochrane diagnostic test accuracy reviews. Syst Rev. 2013;2:82. [PMID: 24099098] doi:10.1186/2046-4053-2-82
14. **Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group.** Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151:264-9. [PMID: 19622511] doi:10.7326/0003-4819-151-4-200908180-00135
15. The 1988 Bethesda System for reporting cervical/vaginal cytological diagnoses. National Cancer Institute Workshop. JAMA. 1989; 262:931-4. [PMID: 2754794]
16. **Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al; Forum Group Members.** The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA. 2002;287: 2114-9. [PMID: 11966386]
17. **Evans DM, Hudson EA, Brown CL, Boddington MM, Hughes HE, Mackenzie EF, et al.** Terminology in gynaecological cytopathology: report of the Working Party of the British Society for Clinical Cytology. J Clin Pathol. 1986;39:933-44. [PMID: 3760239]
18. **Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J.** The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25. [PMID: 14606960]
19. **Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al; QUADAS-2 Group.** QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529-36. [PMID: 22007046] doi:10.7326/0003-4819-155-8-201110180-00009
20. **Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA.** A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics. 2007;8:239-51. [PMID: 16698768]
21. **Harbord RM, Whiting P.** metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. Stata J. 2009;9: 211-29.
22. **Deeks JJ, Bossuyt PM, Gatsonis C.** Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 1.0.0. London: The Cochrane Collaboration; 2013.
23. **Takwoingi Y, Leeflang MM, Deeks JJ.** Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med. 2013;158:544-54. [PMID: 23546566] doi:10.7326/0003-4819-158-7-201304020-00006
24. **Deeks JJ, Macaskill P, Irwig L.** The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005;58: 882-93. [PMID: 16085191]
25. **Katki HA, Schiffman M, Castle PE, Fetterman B, Poitras NE, Lorey T, et al.** Benchmarking CIN 3+ risk as the basis for incorporating HPV and Pap cotesting into cervical screening and management guidelines. J Low Genit Tract Dis. 2013;17:S28-35. [PMID: 23519302] doi:10.1097/LGT.0b013e318285423c
26. **Wentzensen N, Fetterman B, Castle PE, Schiffman M, Wood SN, Stiemerling E, et al.** p16/Ki-67 dual stain cytology for detection of cervical precancer in HPV-positive women. J Natl Cancer Inst. 2015; 107:djv257. [PMID: 26376685] doi:10.1093/jnci/djv257
27. **Arbyn M, Roelens J, Martin-Hirsch P, Leeson S, Wentzensen N.** Use of HC2 to triage women with borderline and mild dyskaryosis in the UK [Editorial]. Br J Cancer. 2011;105:877-80. [PMID: 21952649] doi:10.1038/bjc.2011.351
28. **Belinson JL, Wu R, Belinson SE, Qu X, Yang B, Du H, et al.** A population-based clinical trial comparing endocervical high-risk HPV testing using Hybrid Capture 2 and Cervista from the SHENCCAST II Study. Am J Clin Pathol. 2011;135:790-5. [PMID: 21502436] doi:10.1309/AJCPKA6ATAPBZ6JQ
29. **Cuzick J, Thomas Cox J, Zhang G, Einstein MH, Stoler M, Trupin S, et al.** Human papillomavirus testing for triage of women with low-grade squamous intraepithelial lesions. Int J Cancer. 2013;132:959-66. [PMID: 22806936] doi:10.1002/ijc.27723
30. **Einstein MH, Martens MG, Garcia FA, Ferris DG, Mitchell AL, Day SP, et al.** Clinical validation of the Cervista HPV HR and 16/18 genotyping tests for use in women with ASC-US cytology. Gynecol Oncol. 2010;118:116-22. [PMID: 20488510] doi:10.1016/j.ygyno.2010.04.013
31. **Fröberg M, Johansson B, Hjerpe A, Andersson S.** Human papillomavirus 'reflex' testing as a screening method in cases of minor cytological abnormalities. Br J Cancer. 2008;99:563-8. [PMID: 18682715] doi:10.1038/sj.bjc.6604504
32. **Gage JC, Schiffman M, Solomon D, Wheeler CM, Gravitt PE, Castle PE, et al.** Risk of precancer determined by HPV genotype combinations in women with minor cytologic abnormalities. Cancer Epidemiol Biomarkers Prev. 2013;22:1095-101. [PMID: 23603204] doi:10.1158/1055-9965.EPI-12-1455
33. **Guo M, Lin CY, Gong Y, Cogdell DE, Zhang W, Lin E, et al.** Human papillomavirus genotyping for the eight oncogenic types can improve specificity of HPV testing in women with mildly abnormal Pap results. Mod Pathol. 2008;21:1037-43. [PMID: 18500260] doi:10.1038/modpathol.2008.93
34. **Halfon P, Benmoura D, Agostini A, Khiri H, Penaranda G, Martineau A, et al.** Evaluation of the clinical performance of the Abbott RealTime High-Risk HPV for carcinogenic HPV detection. J Clin Virol. 2010;48:246-50. [PMID: 20541966] doi:10.1016/j.jcv.2010.05.008
35. **Huang S, Erickson B, Tang N, Mak WB, Salituro J, Robinson J, et al.** Clinical performance of Abbott RealTime High Risk HPV test for detection of high-grade cervical intraepithelial neoplasia in women with abnormal cytology. J Clin Virol. 2009;45 Suppl 1:S19-23. [PMID: 19651364] doi:10.1016/S1386-6532(09)70004-6
36. **Lapierre SG, Sauthier P, Mayrand MH, Dufresne S, Petignat P, Provencher D, et al.** Human papillomavirus (HPV) DNA triage of women with atypical squamous cells of undetermined significance with cobas 4800 HPV and Hybrid Capture 2 tests for detection of high-grade lesions of the uterine cervix. J Clin Microbiol. 2012;50: 1240-4. [PMID: 22301023] doi:10.1128/JCM.06656-11
37. **Wong AK, Chan RC, Nichols WS, Bose S.** Invader human papillomavirus (HPV) type 16 and 18 assays as adjuncts to HPV screening of cervical Papanicolaou smears with atypical squamous cells of undetermined significance. Cancer. 2009;115:823-32. [PMID: 19127558] doi:10.1002/cncr.24091
38. **Monsonego J, Pollini G, Evrard MJ, Sednaoui P, Monfort L, Quinzat D, et al.** Linear array genotyping and Hybrid Capture II assay in detecting human papillomavirus genotypes in women referred for colposcopy due to abnormal Papanicolaou smear. Int J STD AIDS. 2008;19:385-92. [PMID: 18595876] doi:10.1258/ijsa.2007.007259
39. **Oliveira A, Verdasca N, Pista Â.** Use of the NucliSENS EasyQ HPV assay in the management of cervical intraepithelial neoplasia. J Med Virol. 2013;85:1235-41. [PMID: 23918542] doi:10.1002/jmv.23590
40. **Depuydt CE, Makar AP, Ruymbeke MJ, Benoy IH, Vereecken AJ, Bogers JJ.** BD-ProExC as adjunct molecular marker for improved detection of CIN2+ after HPV primary screening. Cancer Epidemiol Biomarkers Prev. 2011;20:628-37. [PMID: 21297040] doi:10.1158/1055-9965.EPI-10-0818
41. **Donà MG, Vocaturo A, Giuliani M, Ronchetti L, Rollo F, Pescarmona E, et al.** p16/Ki-67 dual staining in cervico-vaginal cytology: correlation with histology, human papillomavirus detection and genotyping in women undergoing colposcopy. Gynecol Oncol. 2012;126:198-202. [PMID: 22588180] doi:10.1016/j.ygyno.2012.05.004
42. **Spathis A, Kottaridi C, Chranioti A, Meristoudis C, Chrelias C, Panayiotides IG, et al.** mRNA and DNA detection of human papillomaviruses in women of all ages attending two colposcopy clinics. PLoS One. 2012;7:e49205. [PMID: 23166611] doi:10.1371/journal.pone.0049205
43. **Stoler MH, Wright TC Jr, Sharma A, Apple R, Gutekunst K, Wright TL; ATHENA (Addressing THE Need for Advanced HPV Diagnostics) HPV Study Group.** High-risk human papillomavirus testing in women with ASC-US cytology: results from the ATHENA HPV

study. Am J Clin Pathol. 2011;135:468-75. [PMID: 21350104] doi:10.1309/AJCPZ5JY6FCVNMOT

44. **Szarewski A, Ambroisine L, Cadman L, Austin J, Ho L, Terry G, et al.** Comparison of predictors for high-grade cervical intraepithelial neoplasia in women with abnormal smears. Cancer Epidemiol Biomarkers Prev. 2008;17:3033-42. [PMID: 18957520] doi:10.1158/1055-9965.EPI-08-0508

45. **Szarewski A, Mesher D, Cadman L, Austin J, Ashdown-Barr L, Ho L, et al.** Comparison of seven tests for high-grade cervical intraepithelial neoplasia in women with abnormal smears: the Predictors 2 study. J Clin Microbiol. 2012;50:1867-73. [PMID: 22422852] doi:10.1128/JCM.00181-12

46. **Wentzensen N, Schwartz L, Zuna RE, Smith K, Mathews C, Gold MA, et al.** Performance of p16/Ki-67 immunostaining to detect cervical cancer precursors in a colposcopy referral population. Clin Cancer Res. 2012;18:4154-62. [PMID: 22675168] doi:10.1158/1078-0432.CCR-12-0270

47. **Wong OG, Lo CK, Szeto E, Cheung AN.** Efficacy of Abbott RealTime High Risk HPV test in evaluation of atypical squamous cells of undetermined significance from an Asian screening population. J Clin Virol. 2011;51:136-8. [PMID: 21530385] doi:10.1016/j.jcv.2011.03.010

48. **Lin CQ, Cui JF, Zhang X, Pan QJ, Chen W, Qiao YL.** Human papillomavirus genotyping to predict the risk of cervical precancerous lesions or cancer in women with minor abnormal cytology in China. Acta Cytol. 2015;59:405-11. [PMID: 26565687] doi:10.1159/000441290

49. **Iwata T, Hasegawa T, Ochiai K, Takizawa K, Umezawa S, Kuramoto H, et al.** Human papillomavirus test for triage of Japanese women with low-grade squamous intraepithelial lesions. Reprod Sci. 2015;22:1509-15. [PMID: 26092278] doi:10.1177/1933719115589408

50. **Nakamura Y, Matsumoto K, Satoh T, Nishide K, Nozue A, Shimabukuro K, et al.** HPV genotyping for triage of women with abnormal cervical cancer screening results: a multicenter prospective study. Int J Clin Oncol. 2015;20:974-81. [PMID: 25652908] doi:10.1007/s10147-015-0789-4

51. **Khunamornpong S, Settakorn J, Sukpan K, Srisomboon J, Intaraphet S, Siriaunkgul S.** High performance of combined HPV testing and genotyping for HPV16/18/52/58 in triaging women with minor cervical cytological abnormalities in northern Thailand. J Med Virol. 2016;88:135-43. [PMID: 26129775] doi:10.1002/jmv.24290

52. **Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J.** Chapter 9: clinical applications of HPV testing: a summary of meta-analyses. Vaccine. 2006;24 Suppl 3:S78-89. [PMID: 16950021]

53. **Schünemann HJ, Schünemann AH, Oxman AD, Brozek J, Glasziou P, Jaeschke R, et al; GRADE Working Group.** Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ. 2008;336:1106-10. [PMID: 18483053] doi:10.1136/bmj.39500.677199.AE

54. **Huang LW, Lin YH, Pan HS, Seow KM, Lin CY.** Human papillomavirus genotyping as a predictor of high-grade cervical dysplasia in women with mildly cytologic abnormalities: a two-year follow-up report. Diagn Cytopathol. 2012;40:673-7. [PMID: 21442773] doi:10.1002/dc.21591

***ANNALS* AWARDS**

Personae Photography Prize: *Annals* awards a $500 prize for the best photograph submitted each year. Personae photographs are pictures that catch people in the context of their lives and that capture personality. Visit www.annals.org/aim/pages/personae-photography for more information.

Ad Libitum Poetry Prize: All poems published within 1 calendar year are automatically entered into our Poetry Prize Contest. The winning poem is selected by a panel led by Dr. Michael LaCombe and 2 or 3 external judges. The prize for the winning poem is $500. Visit www.annals.org/aim/pages/poetry-prize for more information.

Junior Investigator Awards: *Annals* and the American College of Physicians recognize excellence among internal medicine trainees and junior investigators with annual awards for original research and scholarly review articles published in *Annals*. Visit www.annals.org/aim/pages/junior-investigator-awards for more information.

Current Author Addresses: Drs. Arbyn and Verdoodt and Ms. Xu: Unit of Cancer Epidemiology/Belgian Cancer Centre, Scientific Institute of Public Health, J Wytsmanstreet 14, B1050 Brussels, Belgium.
Dr. Cuzick: Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, United Kingdom.
Dr. Belinson: Cleveland Clinic, Women's Health Institute, 9500 Euclid Avenue, A81, Cleveland, OH 44118.
Drs. Wentzensen and Gage: Division of Cancer Epidemiology and Genetics, National Cancer Institute, 609 Medical Center Drive, Bethesda, MD 20892.
Dr. Khan: Department of Obstetrics and Gynecology, Kaiser Permanente Northern California, 2500 Merced Street, San Leandro, CA 94577.

Author Contributions: Conception and design: M. Arbyn, A. Szarewski, J.C. Gage, M.J. Khan.
Analysis and interpretation of the data: M. Arbyn, L. Xu, F. Verdoodt, J. Cuzick, J.L. Belinson, N. Wentzensen, J.C. Gage, M.J. Khan.
Drafting of the article: M. Arbyn, L. Xu, F. Verdoodt, J. Cuzick, J.C. Gage, M.J. Khan.
Critical revision of the article for important intellectual content: L. Xu, F. Verdoodt, J. Cuzick, J.L. Belinson, N. Wentzensen, J.C. Gage, M.J. Khan.
Final approval of the article: M. Arbyn, L. Xu, F. Verdoodt, J. Cuzick, J.L. Belinson, N. Wentzensen, J.C. Gage, M.J. Khan.
Provision of study materials or patients: M. Arbyn, L. Xu, A. Szarewski, J.C. Gage, M.J. Khan.
Statistical expertise: M. Arbyn, L. Xu, F. Verdoodt, J. Cuzick, N. Wentzensen.
Obtaining of funding: M. Arbyn.
Administrative, technical, or logistic support: M. Arbyn.
Collection and assembly of data: M. Arbyn, L. Xu, F. Verdoodt, A. Szarewski, J.L. Belinson, N. Wentzensen, J.C. Gage, M.J. Khan.

18. References

1. IARC Monograph Working Group. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol 90: Human Papillomaviruses.Editors: Cogliano V, Baan R, Straif K, et al. Lyon: IARCPress; 2007.

2. Campo MS. Papillomavirus research: from natural history to vaccines and beyond. In: Campo MS, editor.Glasgow: Caister Academic Press; 2005. p. 1-424.

3. Bernard HU, Burk RD, Chen Z, et al. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010;**401**: 70-9.

4. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. *Virology* 2004;**324**: 17-27.

5. Bernard HU. The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *J Clin Virol* 2005;**32 Suppl 1**: S1-S6.

6. Van Ranst M, Kaplan JB, Burk RD. Phylogenetic classification of human papillomaviruses: correlation with clinical manifestations. *J Gen Virol* 1992;**73 (Pt 10)**: 2653-60.

7. Richart RM. Cervical intraepithelial neoplasia. *Pathol Annu* 1973;**8**: 301-23.

8. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002;**287**: 2114-9.

9. Bosch FX, Lorincz AT, Munoz N, Meijer CJLM, Shah KV. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002;**55**: 244-65.

10. Castellsague X, Diaz M, de Sanjose S, et al. Worldwide human papillomavirus etiology of cervical adenocarcinoma and its cofactors: implications for screening and prevention. *J Natl Cancer Inst* 2006;**98**: 303-15.

11. Burghardt E. Latest aspects of precancerous lesions in squamous and columnar epithelium of the cervix. *Int J Gynecol Obstet* 1970;**8**: 573-80.

12. Burghardt E, Pickel H, Girardi F. Colposcopy Cervical Pathology. 3rd revised and enlarged edition ed. Stuttgart: Georg Thieme Verlag; 1998.

13. Boon ME, Suurmeijer AJH. The Pap Smear. second edition ed. Coulomb Press Leyden; 1993.

14. Reich O, Regauer S, McCluggage WG, Bergeron C, Redman C. Defining the Cervical Transformation Zone and Squamocolumnar Junction: Can We Reach a Common Colposcopic and Histologic Definition? *Int J Gynecol Pathol* 2017.

15. Herfs M, Yamamoto Y, Laury A, et al. A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proc Natl Acad Sci U S A* 2012;**109**: 10516-21.

16. Herfs M, Parra-Herran C, Howitt BE, et al. Cervical Squamocolumnar Junction-specific Markers Define Distinct, Clinically Relevant Subsets of Low-grade Squamous Intraepithelial Lesions. *Am J Surg Pathol* 2013;**37**: 1311-8.

17. Bouvard V, Baan R, Straif K, et al. A review of human carcinogens--Part B: biological agents. *Lancet Oncol* 2009;**10**: 321-2.

18. Schiffman M, Clifford G, Buonaguro FM. Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline. *Infect Agent Cancer* 2009;**4**: 1-8.

19. Arbyn M, Tommasino M, Depuydt C, Dillner J. Are twenty human papillomavirus types causing cervical cancer? *J Pathol* 2014;**234**: 431-5.

20. Schiffman MA, Herrero R, Desalle R, et al. The carcinogenicity of human papillomavirus types reflects viral evolution. *Virology* 2005;**337**: 76-84.

21. Munoz N, Bosch FX, Castellsague X, et al. Against which human papillomavirus types shall we vaccinate and screen? The international perspective. *Int J Cancer* 2004;**111**: 278-85.

22. Cogliano V, Baan R, Straif K, et al. Carcinogenicity of human papillomaviruses. *Lancet Oncol* 2005;**6**: 204.

23. zur Hausen H. Papillomaviruses in the causation of human cancers - a brief historical account. *Virology* 2009;**384**: 260-5.

24. Werness BA, Levine AJ, Howley PM. Association of human papillomavirus types 16 and 18 E6 proteins with p53. *Science* 1990;**248**: 76-9.

25. von Knebel-Doeberitz M. Papillomaviruses in human disease: Part II. Molecular biology and immunology of papillomavirus infections and carcinogenesis. *Eur J Med* 1992;**1**: 485-91.

26. von Knebel-Doeberitz M. Papillomaviruses in human disease: Part I. Pathogenesis and epidemiology of human papillomavirus infections. *Eur J Med* 1992;**1**: 415-23.

27. von Knebel DM, Rittmuller C, Aengeneyndt F, Jansen-D P. Reversible repression of papillomavirus oncogene expression in cervical carcinoma cells:

consequences for the phenotype and e6- p53 and e7-prb interactions. *J Virol* 1994;**68**: 2811-21.

28. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Biological Agents: A Review of Human Carcinogens. *IARC Monogr Eval Carcinog Risks Hum* 2012;**100B**: 1-475.

29. Vinokurova S, Wentzensen N, Kraus I, et al. Type-dependent integration frequency of human papillomavirus genomes in cervical lesions. *Cancer Res* 2008;**68**: 307-13.

30. von Knebel-Doeberitz M. Papillomaviruses in human disease: Part II. Molecular biology and immunology of papillomavirus infections and carcinogenesis. *Eur J Med* 1992;**1**: 485-91.

31. Snijders PJ, Steenbergen RD, Heideman DA, Meijer CJLM. HPV-mediated cervical carcinogenesis: concepts and clinical implications. *J Pathol* 2006;**208**: 152-64.

32. Steenbergen RD, Snijders PJ, Heideman DA, Meijer CJLM. Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions. *Nat Rev Cancer* 2014;**14**: 395-405.

33. Stevanovic S, Pasetto A, Helman SR, et al. Landscape of immunogenic tumor antigens in successful immunotherapy of virally induced epithelial cancer. *Science* 2017;**356**: 200-5.

34. Cancer Genome Atlas Research Network. Integrated genomic and molecular characterization of cervical cancer. *Nature* 2017;**543**: 378-84.

35. Macfarlane LA, Murphy PR. MicroRNA: Biogenesis, Function and Role in Cancer. *Curr Genomics* 2010;**11**: 537-61.

36. Lacey CJ, Lowndes CM, Shah KV. Chapter 4: Burden and management of non-cancerous HPV-related conditions: HPV-6/11 disease. *Vaccine* 2006;**24**: 35-41.

37. Winer RL, Lee SK, Hughes JP, et al. Genital Human Papillomavirus infection: Incidence and risk factors in a cohort of female university students. *J Epidemiol* 2003;**157**: 218-26.

38. Winer RL, Feng Q, Hughes JP, et al. Risk of female human papillomavirus acquisition associated with first male sex partner. *J Infect Dis* 2008;**197**: 279-82.

39. de Sanjose S, Diaz M, Castellsague X, et al. Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis. *Lancet Infect Dis* 2007;**7**: 453-9.

40. Ho GYF, Bierman R, Beardsley I, Chang CJ, Burk RD. Natural history of cervicovaginal papillomavirus infection in young women. *N Engl J Med* 1998;**338**: 423-8.

41. 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**: 425-34.

42. Ostor AG. Natural history of cervical intraepithelial neoplasia: a critical review. *Int J Gynecol Pathol* 1993;**12**: 186-92.

43. IARC. Cervix Cancer Screening. IARC Handbooks of Cancer Prevention. Vol. 10. Lyon: IARCPress; 2005.

44. Bray F, Loos AH, McCarron P, et al. Trends in cervical squamous cell carcinoma incidence in 13 European countries: changing risk and the effects of screening. *Cancer Epidemiol Biomarkers Prev* 2005;**14**: 677-86.

45. Arbyn M, Raifu AO, Antoine J. Trends of cervical cancer mortality in Europe. *IPH/EPI-REPORTS* 2009;**D/2009/2505/41**: 1-54.

46. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;**136**: E359-E386.

47. Arbyn M, Castellsagué X, de Sanjosé S, et al. Worldwide burden of cervical cancer in 2008. *Ann Oncol* 2011;**22**: 2675-86.

48. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;**127**: 2893-917.

49. Smith HO, Tiffany MF, Qualls CR, Key CR. The rising incidence of adenocarcinoma relative to squamous cell carcinoma of the uterine cervix in the United States--a 24-year population-based study. *Gynecol Oncol* 2000;**78**: 97-105.

50. Bray F, Carstensen B, Moller H, et al. Incidence trends of adenocarcinoma of the cervix in 13 European countries. *Cancer Epidemiol Biomarkers Prev* 2005;**14**: 2191-9.

51. Yang BH, Bray FI, Parkin DM, Sellors JW, Zhang Z-F. Cervical cancer as a priority for prevention in different world regions : an evaluation using years of life lost. *Int J Cancer* 2004;**109**: 418-24.

52. Fitzmaurice C, Allen C, Barber RM, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A

Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 2016.

53. Sant M, Allemani C, Santaquilani M, et al. EUROCARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. *Eur J Cancer* 2009;**45**: 931-91.

54. Miller AB. Cervical cancer screening programmes. Managerial guidelines. Geneva: World Health Organisation; 1993.

55. Martin-Hirsch PP, Paraskevaidis E, Bryant A, Dickinson HO, Keep SL. Surgery for cervical intraepithelial neoplasia. *Cochrane Database Syst Rev* 2010; CD001318.

56. Singer A, Monaghan JM. Lower Genital Tract Precancer: Colposcopy, Pathology and Treatment. 2nd ed. Oxford: Blackwell Science Ltd; 2000.

57. Hakama M, Miller AB, Day NE. Screening for Cancer of the Uterine Cervix. Hakama M, Miller AB, Day NE, editors. 1-308. 1986. Lyon, International Agency for Research on Cancer. IARC Scientific Publications.

58. Sasieni P, Adams J, Cuzick J. Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. *Br J Cancer* 2003;**89**: 88-93.

59. van den Akker van Marle ME, van Ballegooijen M, Habbema JD. Low risk of cervical cancer during a long period after negative screening in the Netherlands. *Br J Cancer* 2003;**88**: 1054-7.

60. Zappa M, Visioli CB, Ciatto S, et al. Lower protection of cytological screening for adenocarcinomas and shorter protection for younger women: the results of a case-control study in Florence. *Br J Cancer* 2004;**90**: 1784-6.

61. Andrae B, Kemetli L, Sparen P, et al. Screening-preventable cervical cancer risks: evidence from a nationwide audit in Sweden. *J Natl Cancer Inst* 2008;**100**: 622-9.

62. Arbyn M, Rebolj M, de Kok IM, et al. The challenges for organising cervical screening programmes in the 15 old member states of the European Union. *Eur J Cancer* 2009;**45**: 2671-8.

63. Arbyn M, Raifu AO, Weiderpass E, Bray F, Anttila A. Trends of cervical cancer mortality in the member states of the European Union. *Eur J Cancer* 2009;**45**: 2640-8.

64. Anttila A, Nieminen P. Cervical cancer screening programme in Finland with an example on implementing alternative screening methods. *Coll Antropol* 2007;**31 Suppl 2:17-22.**: 17-22.

65. Boyes DA, Morrison B, Knox EG, Draper GJ, Miller AB. A cohort study of cervical cancer screening in British Columbia. *Clin Invest Med* 1982;**5**: 1-29.

66. Miller A, Sankaranarayanan R, Bosch FX, WHO. Cervical Cancer Screening in Developing Countries. Geneva: World Health Organisation; 2002.

67. Bruni L, Diaz M, Castellsagué X, et al. Cervical HPV prevalence in five continents: meta-analysis on one million women with normal cytology. *J Infect Dis* 2010;**202**: 1789-99.

68. Arbyn M, Sasieni P, Meijer CJLM, et al. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. *Vaccine* 2006;**24 (SUPPL. 3)**: S78-S89.

69. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding HPV testing in secondary prevention of cervical cancer. *Vaccine* 2012;**30 Suppl 5**: F88-F99.

70. Cuzick J, Arbyn M, Ronco G, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. *Vaccine* 2008;**26 Suppl 10**: K29-K41.

71. Wright TC, Cox JT, Massad LS, Wilkinson EJ. 2001 Consensus guidelines for the management of women with cervical cytological abnormalities. *JAMA* 2002;**287**: 2120-9.

72. Jordan J, Martin-Hirsch P, Arbyn M, et al. European guidelines for management of abnormal cervical cytology, Part 2. *Cytopathology* 2009;**20**: 5-16.

73. Cox JT. HPV testing: is it useful in triage of minor Pap abnormalities? *J Fam Pract* 1998;**46**: 121-4.

74. Cox JT. Management of women with cervical cytology interpreted as ASC-US or as ASC-H. *Clin Obstet Gynecol* 2005;**48**: 160-77.

75. Sawaya GF. A 21-year-old woman with atypical squamous cells of undetermined significance. *JAMA* 2005;**294**: 2210-8.

76. Soutter WP. Management of women with mild dyskaryosis. Immediate referral to colposcopy is safer. *BMJ* 1994;**309**: 591-2.

77. Stoler MH, Schiffman MA. Interobserver reproducibility of cervical cytologic and histologic interpretations. *JAMA* 2001;**285**: 1500-5.

78. Lundberg GD, National Cancer Institute. The 1988 Bethesda System for Reporting Cervical/Vaginal Cytologic Diagnoses. *JAMA* 1989;**262**: 931-4.

79. Dudding N, Sutton J. BSCC Terminology Conference, Koilocytosis and Mild Dyskaryosis. *Cytopathology* 2002;**13**: 379-81.

80. Robertson JH, Woodend BE, Crozier EH, Hutchinson J. Risk of cervical cancer associated with mild dyskaryosis. *BMJ* 1988;**297**: 18-21.

81. Coleman D, Day N, Douglas G, et al. European Guidelines for Quality Assurance in Cervical Cancer Screening. Europe against cancer programme. *Eur J Cancer* 1993;**29A Suppl 4**: S1-S38.

82. Flannelly G, Anderson D, Kitchener HC, et al. Management of women with mild and moderate cervical dyskaryosis. *BMJ* 1994;**308**: 1399-403.

83. Richart RM. Causes and management of cervical intraepithelial neoplasia. *Cancer* 1987;**60**: 1951-9.

84. Noumoff JS. Atypia in cervical cytology as a risk factor for intraepithelial neoplasia. *Am J Obstet Gynecol* 1987;**156**: 628-31.

85. Richart RM, Wright TC. Controversies in the management of low-grade cervical intraepithelial neoplasia. *Cancer* 1993;**71**: 1413-21.

86. Cox JT, Lorincz AT, Schiffman MH, et al. Human papillomavirus testing by hybrid capture appears to be useful in triaging women with a cytologic diagnosis of atypical squamous cells of undetermined significance. *Am J Obstet Gynecol* 1995;**172**: 946-54.

87. Wright TC, Sun XW, Koulos J. Comparison of management algorithms for the evaluation of women with low-grade cytologic abnormalities. *Obstet Gynecol* 1995;**85**: 202-10.

88. Kinney WK, Manos MM, Hurley LB, Ransley JE. Where's the high grade cervical neoplasia? The importance of minimally abnormal Papanicolaou diagnoses. *Obstet Gynecol* 1998;**91**: 973-6.

89. Arbyn M, Kyrgiou M, Simoens C, et al. Peri-natal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: a meta-analysis. *BMJ* 2008;**337**: a1284.

90. Kyrgiou M, Koliopoulos G, Martin-Hirsch P, et al. Obstetric outcomes after conservative treatment for intra-epithelial or early invasive cervical lesions: a systematic review and meta-analysis of the literature. *Lancet* 2006;**367**: 489-98.

91. zur Hausen H. Molecular pathogenesis of cancer of the cervix and its causation by specific human papillomavirus types. *Curr Top Microbiol Immunol* 1994;**186**: 131-56.

92. Walboomers JM, Jacobs MV, Manos M, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999;**189**: 12-9.

93. Arbyn M, Buntinx F, Van Ranst M, et al. Virologic versus cytologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. *J Natl Cancer Inst* 2004;**96**: 280-93.

94. Arbyn M, Roelens J, Simoens C, et al. Human papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions. *Cochrane Database Syst Rev* 2013;**3**: 1-201.

95. Moscicki AB, Schiffman M, Burchell A, et al. Updating the natural history of human papillomavirus and anogenital cancers. *Vaccine* 2012;**30 Suppl 5**: F24-F33.

96. Paraskevaidis E, Arbyn M, Diakomanolis E, et al. The role of HPV DNA testing in the follow-up period after treatment for CIN: a systematic review of the literature. *Cancer Treat Rev* 2004;**30**: 205-11.

97. Alonso I, Torne A, Puig-Tintore LM, et al. Pre- and post-conization high-risk HPV testing predicts residual/recurrent disease in patients treated for CIN 2-3. *Gynecol Oncol* 2006;**103**: 631-6.

98. Stern PL, van der Burg SH, Hampson IN, et al. Therapy of human papillomavirus-related disease. *Vaccine* 2012;**30 Suppl 5**: F71-F82.

99. Melnikow J, McGahan C, Sawaya GF, Ehlen T, Coldman A. Cervical Intraepithelial Neoplasia Outcomes After Treatment: Long-term Follow-up From the British Columbia Cohort Study. *J Natl Cancer Inst* 2009;**110**: 721-8.

100. Soutter WP, Sasieni P, Panoskaltsis T. Long-term risk of invasive cervical cancer after treatment of squamous cervical intraepithelial neoplasia. *Int J Cancer* 2005;**118**: 2048-55.

101. Kalliala I, Anttila A, Pukkala E, Nieminen P. Risk of cervical and other cancers after treatment of cervical intraepithelial neoplasia: retrospective cohort study. *BMJ* 2005;**331**: 1183-5.

102. Strander B, Hallgren J, Sparen P. Effect of ageing on cervical or vaginal cancer in Swedish women previously treated for cervical intraepithelial neoplasia grade 3: population based cohort study of long term incidence and mortality. *BMJ* 2014;**348**: f7361.

103. Arbyn M, Kyrgiou M, Gondry J, Petry KU, Paraskevaidis E. Long term outcomes for women treated for cervical precancer. *BMJ* 2014;**348**: f7700.

104. Arbyn M, Sasieni P, Meijer CJ, et al. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. *Vaccine* 2006;**24 (SUPPL. 3)**: S78-S89.

105. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding HPV testing in secondary prevention of cervical cancer. *Vaccine* 2012;**30 Suppl 5**: F88-F99.

106. Bosch FX, Lorincz AT, Munoz N, Meijer CJ, Shah KV. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002;**55**: 244-65.

107. Koliopoulos G, Martin-Hirsch P, Paraskevaidis E, et al. HPV testing versus cervical cytology for screening for cancer of the uterine cervix (Protocol). *Cochrane Database Syst Rev* 2010;**4**: CD-0044709.

108. Koliopoulos G, Arbyn M, Martin-Hirsch P, et al. Diagnostic accuracy of human papillomavirus testing in primary cervical screening: a systematic review and meta-analysis of non randomised studies. *Gynecol Oncol* 2007;**104**: 232-46.

109. European Commission D-GfHaFS, Anttila A, Arbyn M, De Vuyst H, Dillner J, Dillner L, et al. European Guidelines for Quality Assurance in Cervical Cancer Screening. Second Edition : Supplements. Luxembourg: Publications Office of the European Union; 2015.

110. von Karsa L, Arbyn M, De Vuyst H, et al. European guidelines for quality assurance in cervical cancer screening. Summary of the supplements on HPV screening and vaccination. *Papillomavir Res* 2015;**1**: 22-31.

111. Arbyn M, Haelens A, Desomer A, et al. Cervical cancer screening program and Human Papillomavirus (HPV) testing, part II: Update on HPV primary screening. *KCE reports* 2015;**238**: 1-282.

112. Verdoodt F, Jentschke M, Hillemanns P, et al. Reaching women who do not participate in the regular cervical cancer screening program by offering self-sampling kits: A systematic review and meta-analysis of randomised trials. *Eur J Cancer* 2015;**51**: 2375-85.

113. Glasziou P, Irwig L, Bain C, Colditz G. Systematic Reviews in Health Care: a Practical Guide. 1 ed. Cambridge: The Press Syndicate of the University of Cambridge; 2001.

114. Rothman KJ, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia: Lippincot-Raven; 1998.

115. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].Editor: The Cochrane Collaboration. 2011.

116. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;**151**: 264-9, W64.

117. Diagnostic Test Accuracy Working Group. Handbook for DTA Reviews. 2010.

118. Arbyn M, Dillner J, Van Ranst M, et al. Re: Have we resolved how to triage equivocal cervical cytology? *J Natl Cancer Inst* 2004;**96**: 1401-2.

119. Arbyn M, Paraskevaidis E, Martin-Hirsch P, Prendiville W, Dillner J. Clinical utility of HPV DNA detection: triage of minor cervical lesions, follow-up of women treated for high-grade CIN. An update of pooled evidence. *Gynecol Oncol* 2005;**99 (Suppl 3)**: 7-11.

120. Arbyn M, Roelens J, Cuschieri K, et al. The APTIMA HPV assay versus the Hybrid Capture II test in triage of women with ASC-US or LSIL cervical cytology: a meta-analysis of the diagnostic accuracy. *Int J Cancer* 2013;**132**: 101-8.

121. Roelens J, Reuschenbach M, von Knebel-Doeberitz M, et al. p16INK4a immunocytochemistry versus HPV testing for triage of women with minor cytological abnormalities: A systematic review and meta-analysis. *Cancer* 2012;**120**: 294-307.

122. Verdoodt F, Szarewski A, Halfon P, Cuschieri K, Arbyn M. Triage of women with minor abnormal cervical cytology: Meta-analysis of the accuracy of an assay targeting messenger ribonucleic acid of 5 high-risk human papillomavirus types. *Cancer Cytopathol* 2013;**121**: 675-87.

123. Arbyn M, Martin-Hirsch P, Buntinx F, et al. Triage of women with equivocal or low-grade cervical cytology results. A meta-analysis of the HPV test positivity rate. *J Cell Mol Med* 2009;**13**: 648-59.

124. Herbert A, Arbyn M, Bergeron C. Why CIN3 and CIN2 should be distinguished on histological reports. *Cytopathology* 2008;**19**: 63-4.

125. Ismail SM, Colclough AB, Dinnen JS, et al. Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia. *BMJ* 1989;**298**: 707-10.

126. Holowaty P, Miller AB, Rohan T, To T. Natural History of Dysplasia of the Uterine Cervix. *J Natl Cancer Inst* 1999;**91**: 252-8.

127. Sherman ME, Schiffman MA, Cox JT. Effects of age and human papilloma viral load on colposcopy triage: data from the randomised atypical squamous cells of undetermined significance/low-grade intraepithelial lesion triage study (ALTS). *J Natl Cancer Inst* 2002;**94**: 102-7.

128. Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS : a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**: 1-13.

129. 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**: 529-36.

130. Higgins JP, Green S. Systematic Reviews of Interventions. Chichester (UK): Wiley-Blackwell; 2008.

131. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**: 239-51.

132. Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. *The Stata Journal* 2009;**9**: 197-210.

133. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2002;**20**: 2865-84.

134. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006;**59**: 1331-2.

135. Harbord RM, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *The Stata Journal* 2009;**9**: 211-29.

136. midas: A program for Meta-analytical Integration of Diagnostic Accuracy Studies in Stata. [computer program]. Division of Nuclear Medicine, Department of Radiology, University of Michigan Medical School, Ann Arbor, Michigan, USA.: 2012.

137. Deeks JJ, Bossuyt PM, Gatsonis C. **Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0** .Editors: Deeks JJ, Bossuyt PM, Gatsonis C. The Cochrane Collaboration; 2013.

138. Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Ann Intern Med* 2013;**158**: 544-54.

139. Nyaga VN, Arbyn M, Aerts M. Metaprop: a Stata command to perform meta-analysis of binomial data. *Arch Public Health* 2014;**72**: 39.

140. Harris R, Bradburn M, Deeks J, et al. metan: fixed- and random-effects meta-analysis. *Stata Journal* 2008;**8**: 3-28.

141. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;**10**: 101-29.

142. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**: 557-60.

143. Harbord R, Higgins JP. Meta-regression in Stata. *Stata J* 2013;**8**: 493-519.

144. Arbyn M, Vanden Broeck D, Benoy I, et al. Surveillance of effects of HPV vaccination in Belgium. *Cancer Epidemiol* 2016;**41**: 152-8.

145. Arbyn M, Xu L, Verdoodt F, et al. Genotyping for human papillomavirus types 16 and 18 in women with minor cervical lesions: a systematic review and meta-analysis. *Ann Intern Med* 2017;**166**: 118-27.

146. Arbyn M, Bryant A, Martin-Hirsch PL, et al. Prophylactic vaccination against human papillomaviruses to prevent cervical cancer and its precursors (Protocol). *Cochrane Database Syst Rev* 2013;**12**: 1-15.

147. Kyrgiou M, Stasinou SM, Arbyn M, et al. Management of low-grade squamous intra-epithelial lesions of the uterine cervix: repeat cytology versus immediate referral to colposcopy (review). *Cochrane Database Syst Rev* 2017; accepted-for publication.

148. Kyrgiou M, Mitra A, Arbyn M, et al. Fertility and early pregnancy outcomes after conservative treatment for cervical intraepithelial lesions. *Cochrane Database Syst Rev* 2015;**9**: CD008478-*.

149. Arbyn M, Fisher A, Martin-Hirsch P, Zhao FH, Qiao YL. Low cost versus other screening tests to detect cervical cancer or precancer in developing countries (Protocol). *The Cochrane Library* 2012.

150. Meijer CJLM, Castle PE, Hesselink AT, 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**: 516-20.

151. Wong OG, Lo CK, Szeto E, Cheung AN. Efficacy of Abbott RealTime High Risk HPV test in evaluation of atypical squamous cells of undetermined significance from an Asian screening population. *J Clin Virol* 2011;**51**: 136-8.

152. Denton KJ, Bergeron C, Klement P, et al. The Sensitivity and Specificity of p16INK4a Cytology vs HPV Testing for Detecting High-Grade Cervical Disease in the Triage of ASC-US and LSIL Pap Cytology Results. *Am J Clin Pathol* 2010;**134**: 12-21.

153. Trope A, Jonassen CM, Sjoborg KD, et al. Role of high-risk human papillomavirus (HPV) mRNA testing in the prediction of residual disease after conisation for high-grade cervical intraepithelial neoplasia. *Gynecol Oncol* 2011;**123**: 157-62.

154. Ryu A, Nam K, Kwak J, Kim J, Jeon S. Early human papillomavirus testing predicts residual/ recurrent disease after LEEP. *J Gynecol Oncol* 2012;**23**: 217-25.

155. Torne A, Fuste P, Rodriguez-Carunchio L, et al. Intraoperative post-conisation human papillomavirus testing for early detection of treatment failure in patients with cervical intraepithelial neoplasia: a pilot study. *BJOG* 2013;**120**: 392-9.

156. Chua KL, Hjerpe A. Human papillomavirus analysis as a prognostic marker following conization of the cervix uteri. *Gynecol Oncol* 1997;**66**: 108-13.

157. Heymans J, Benoy IH, Poppe W, Depuydt CE. Type-specific HPV geno-typing improves detection of recurrent high-grade cervical neoplasia after conisation. *Int J Cancer* 2011;**129**: 903-9.

158. Zielinski GD, Rozendaal L, Voorhorst FJ, et al. HPV testing can reduce the number of follow-up visits in women treated for cervical intraepithelial neoplasia grade 3. *Gynecol Oncol* 2003;**91**: 67-73.

159. Sarian LO, Derchain SF, Andrade LA, et al. HPV DNA test and Pap smear in detection of residual and recurrent disease following loop electrosurgical excision procedure of high-grade cervical intraepithelial neoplasia. *Gynecol Oncol* 2004;**94**: 181-6.

160. Kreimer AR, Guido RS, Solomon D, et al. Human papillomavirus testing following loop electrosurgical excision procedure identifies women at risk for posttreatment cervical intraepithelial neoplasia grade 2 or 3 disease. *Cancer Epidemiol Biomarkers Prev* 2006;**15**: 908-14.

161. Verguts J, Bronselaer B, Donders G, et al. Prediction of recurrence after treatment for high-grade cervical intraepithelial neoplasia: the role of human papillomavirus testing and age at conisation. *BJOG* 2006;**113**: 1303-7.

162. Kang WD, Jeong OM, Kim SM, et al. Significance of human papillomavirus genotyping with high-grade cervical intraepithelial neoplasia treated by a loop electrosurgical excision procedure. *Am J Obstet Gynecol* 2010;**203**: 72-6.

163. Smart OC, Sykes P, MacNab H, Jennings L. Testing for high risk human papilloma virus in the initial follow-up of women treated for high-grade squamous intraepithelial lesions. *Aust N Z J Obstet Gynaecol* 2010;**50**: 164-7.

164. Nobbenhuis MA, Meijer CJLM, van den Brule AJ, et al. Addition of high-risk HPV testing improves the current guidelines on follow-up after treatment for cervical intraepithelial neoplasia. *Br J Cancer* 2001;**84**: 796-801.

165. Cecchini S, Carozzi F, Confortini M, Zappa M, Ciatto S. Persistent human papilloma virus infection as an indicator of risk of recurrence of high-grade cervical intraepithelial neoplasia treated by the loop electrosurgical excision procedure. *Tumori* 2004;**90**: 225-8.

166. Fambrini M, Penna C, Pieralli A, et al. PCR detection rates of high risk human papillomavirus DNA in paired self-collected urine and cervical scrapes after

laser CO2 conization for high-grade cervical intraepithelial neoplasia. *Gynecol Oncol* 2008.

167. Aerssens A, Claeys P, Beerens E, et al. Prediction of recurrent disease by cytology and HPV testing after treatment of cervical intraepithelial neoplasia. *Cytopathology* 2009;**20**: 27-35.

168. Bais AG, Eijkemans MJ, Rebolj M, et al. Post-treatment CIN: randomised clinical trial using hrHPV testing for prediction of residual/recurrent disease. *Int J Cancer* 2009;**124**: 889-95.

169. Jeong NH, Lee NW, Kim HJ, Kim T, Lee KW. High-risk human papillomavirus testing for monitoring patients treated for high-grade cervical intraepithelial neoplasia. *J Obstet Gynaecol Res* 2009;**35**: 706-11.

170. Ang C, Mukhopadhyay A, Burnley C, et al. Histological recurrence and depth of loop treatment of the cervix in women of reproductive age: incomplete excision versus adverse pregnancy outcome. *BJOG* 2011;**118**: 685-92.

171. Flannelly G, Bolger B, Fawzi H, De Lopes AB, Monaghan JM. Follow up after LLETZ: could schedules be modified according to risk of recurrence? *BJOG* 2001;**108**: 1025-30.

172. Prato B, Ghelardi A, Gadducci A, et al. Correlation of recurrence rates and times with posttreatment human papillomavirus status in patients treated with loop electrosurgical excision procedure conization for cervical squamous intraepithelial lesions. *Int J Gynecol Cancer* 2008;**18**: 90-4.

173. Strander B, Andersson-Ellstrom A, Milsom I, Sparen P. Long term risk of invasive cancer after treatment for cervical intraepithelial neoplasia grade 3: population based cohort study. *BMJ* 2007;**335**: 1077.

174. Castle PE, Schiffman MA, Herrero R, et al. A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste, Costa Rica. *J Infect Dis* 2005;**191**: 1808-16.

175. Aerssens A, Claeys P, Garcia A, et al. Natural history and clearance of HPV after treatment of precancerous cervical lesions. *Histopathology* 2008;**52**: 381-6.

176. Sarian LO, Derchain SF, Pitta DR, Morais SS, Rabelo-Santos SH. Factors associated with HPV persistence after treatment for high-grade cervical intra-epithelial neoplasia with large loop excision of the transformation zone (LLETZ). *J Clin Virol* 2004;**31**: 270-4.

177. Fambrini M, Penna C, Pieralli A, et al. CO2 laser cylindrical excision or standard re-conization for persistent-recurrent high-grade cervical intraepithelial neoplasia (HG-CIN) in women of fertile age. *Anticancer Res* 2008;**28**: 3871-5.

178. Poljak M, Cuzick J, Kocjan BJ, et al. Nucleic acid tests for the detection of human papillomaviruses. *Vaccine* 2012;**30 (Suppl 5)**: F100-F106.

179. Arbyn M, Sankaranarayanan R, Muwonge R, et al. Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India. *Int J Cancer* 2008;**123**: 153-60.

180. Catarino R, Schafer S, Vassilakos P, Petignat P, Arbyn M. Accuracy of combinations of visual inspection using acetic acid or lugol iodine to detect cervical precancer: a Meta-analysis. *BJOG* 2017.

181. Belinson JL, Qiao YL, Pretorius R, et al. Shanxi province cervical cancer screening study: a cross-sectional comparative trial multiple techniques to detect cervical neoplasia. *Gynecol Oncol* 2001;**83**: 439-44.

182. Arbyn M, de Sanjose S, Saraiya M, et al. EUROGIN 2011 roadmap on prevention and treatment of HPV-related disease. *Int J Cancer* 2012;**131**: 1969-82.

183. Sankaranarayanan R, Nene BM, Dinshaw KA, et al. A cluster randomized controlled trial of visual, cytology and human papillomavirus screening for cancer of the cervix in rural India. *Int J Cancer* 2005;**116**: 617-23.

184. Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in Rural India. *N Engl J Med* 2009;**360**: 1385-94.

185. 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**: 672-82.

186. Ronco G, Segnan N, Giorgi-Rossi P, et al. Human Papillomavirus testing and liquid-based cytology in primary cervical screening: results at recruitment from the New Technologies for Cervical Cancer randomized controlled trial. *J Natl Cancer Inst* 2006;**98**: 765-74.

187. Mayrand MH, Duarte-Franco E, Rodrigues I, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med* 2007;**357**: 1579-88.

188. Rijkaart DC, Berkhof J, Rozendaal L, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol* 2012;**13**: 78-88.

189. Naucler P, Ryd W, Tornberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med* 2007;**357**: 1589-97.

190. Leinonen M, Nieminen P, Kotaniemi-Talonen L, et al. Age-Specific Evaluation of Primary Human Papillomavirus Screening vs Conventional Cytology in a Randomized Setting. *J Natl Cancer Inst* 2009;**101**: 1612-23.

191. Naucler P, Ryd W, Tornberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med* 2007;**357**: 1589-97.

192. 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**: 672-82.

193. Rijkaart DC, Berkhof J, Rozendaal L, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol* 2012;**13**: 78-88.

194. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010;**11**: 249-57.

195. Anttila A, Kotaniemi-Talonen L, Leinonen M, et al. Rate of cervical cancer, severe intraepithelial neoplasia, and adenocarcinoma in situ in primary HPV DNA screening with cytology triage: randomised study within organised screening programme. *BMJ* 2010;**340**: c1804.

196. Leinonen M, Nieminen P, Kotaniemi-Talonen L, et al. Age-Specific Evaluation of Primary Human Papillomavirus Screening vs Conventional Cytology in a Randomized Setting. *J Natl Cancer Inst* 2009;**101**: 1612-23.

197. Leinonen MK, Nieminen P, Lonnberg S, et al. Detection rates of precancerous and cancerous cervical lesions within one screening round of primary human papillomavirus DNA testing: prospective randomised trial in Finland. *BMJ* 2012;**345**: e7789-0.

198. Ogilvie GS, Van Niekerk DJ, Krajden M, et al. A randomized controlled trial of Human Papillomavirus (HPV) testing for cervical cancer screening: trial design and preliminary results (HPV FOCAL Trial). *BMC Cancer* 2010;**10**: 111.

199. Ronco G, Dillner J, Elfstrom KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet* 2014;**383**: 524-32.

200. Castle PE, Stoler MH, Wright TC, Jr., et al. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. *Lancet Oncol* 2011;**12**: 880-90.

201. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Human papillomavirus testing and liquid-based cytology in primary screening of women younger than 35 years: results at recruitment for a randomised controlled trial. *Lancet Oncol* 2006;**7**: 547-55.

202. Carozzi F, Confortini M, Palma PD, et al. Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol* 2008;**9**: 937-45.

203. Carozzi F, Gillio-Tos A, Confortini M, et al. Risk of high-grade cervical intraepithelial neoplasia during follow-up in HPV-positive women according to baseline p16-INK4A results: a prospective analysis of a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol* 2013;**14**: 168-76.

204. Naucler P, Ryd W, Tornberg S, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *J Natl Cancer Inst* 2009;**101**: 88-98.

205. Rijkaart DC, Berkhof J, van Kemenade FJ, et al. Evaluation of 14 triage strategies for HPV DNA-positive women in population-based cervical screening. *Int J Cancer* 2012;**130**: 602-10.

206. Dijkstra M, Van ND, Rijkaart D, et al. Primary hrHPV DNA testing in Cervical Cancer screening: how to manage screen positive women? A POBASCAM Trial sub study. *Cancer Epidemiol Biomarkers Prev* 2013; *.

207. Leinonen MK, Anttila A, Malila N, et al. Type- and age-specific distribution of human papillomavirus in women attending cervical cancer screening in Finland. *Br J Cancer* 2013;**109**: 2941.

208. Castle PE, Stoler MH, Wright TC, Jr., et al. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. *Lancet Oncol* 2011;**12**: 880-90.

209. Carozzi F, Gillio-Tos A, Confortini M, et al. Risk of high-grade cervical intraepithelial neoplasia during follow-up in HPV-positive women according to baseline p16-INK4A results: a prospective analysis of a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol* 2012.

210. Petry KU, Schmidt D, Scherbring S, et al. Triaging Pap cytology negative, HPV positive cervical cancer screening results with p16/Ki-67 Dual-stained cytology. *Gynecol Oncol* 2011;**121**: 505-9.

211. Dillner J, Rebolj M, Birembaut P, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ* 2008;**337**: a1754.

212. Clavel C, Masure M, Bory J-P, et al. Human papillomavirus testing in primary screening for the detection of high-grade cervical lesions: a study of 7932 women. *Br J Cancer* 2001;**89**: 1616-23.

213. Kjaer SK, van den Brule AJ, Paull G, et al. Type specific persistence of high risk human papillomavirus (HPV) as indicator of high grade cervical squamous intraepithelial lesions in young women: population based prospective follow up study. *BMJ* 2002;**325**: 572-d.

214. Petry KU, Menton S, Menton M, et al. Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8466 patients. *Br J Cancer* 2003;**88**: 1570-7.

215. de Sanjose S, Almirall R, Lloveras B, et al. Cervical Human Papillomavirus Infection in the Female Population in Barcelona, Spain. *Sex Transm Dis* 2003;**30**: 788-93.

216. Cuzick J, Szarewski A, Mesher D, et al. Long-term follow-up of cervical abnormalities among women screened by HPV testing and cytology-Results from the Hammersmith study. *Int J Cancer* 2008;**122**: 2294-300.

217. Katki HA, Schiffman M, Castle PE, et al. Benchmarking CIN 3+ Risk as the Basis for Incorporating HPV and Pap Cotesting into Cervical Screening and Management Guidelines. *J Low Genit Tract Dis* 2013;**17**: S28-S35.

218. Sherman ME, Lorincz AT, Scott DR, et al. Baseline Cytology, Human Papillomavirus Testing, and Risk for Cervical Neoplasia: A 10-Year Cohort Analysis. *J Natl Cancer Inst* 2003;**95**: 46-52.

219. Castle PE, Glass AG, Rush BB, et al. Clinical human papillomavirus detection forecasts cervical cancer risk in women over 18 years of follow-up. *J Clin Oncol* 2012;**30**: 3044-50.

220. Katki HA, Kinney WK, Fetterman B, et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. *Lancet Oncol* 2011;**12**: 666-72.

221. Arbyn M, Anttila A, Jordan J, et al. European Guidelines for Quality Assurance in Cervical Cancer Screening. Second Edition - Summary Document. *Ann Oncol* 2010;**21**: 448-58.

222. Advisory Committee on Cancer Prevention. Recommendations on cancer screening in the European Union. Advisory Committee on Cancer Prevention. *Eur J Cancer* 2000;**36**: 1473-8.

223. European Commission. European Guidelines for Quality Assurance in Cervical Cancer Screening. 2nd ed. Editors: Arbyn M, Anttila A, Jordan J, Ronco G,

Schenck U, Segnan N et al. Luxembourg: Office for Official Publications of the European Communities; 2008.

224. European Commission. European guidelines for quality asurance in breast cancer screening and diagnosis.Editors: Perry N, Broeders M, de Wolf C, Törnberg S. Luxembourg: Office for Official Publications of the European Communities; 2006.

225. Anttila A, Ronco G, Lynge E, Fender M, Arbyn M, Baldauf JJ, et al. Chapter 2: Epidemiological Guidelines for Quality Assurance in Cervical Cancer Screening. In: Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, et al., editors. European Guidelines for Quality Assurance in Cervical Cancer Screening. 2nd ed. Luxembourg: Office for Official Publications of the European Communities; 2008. p. 11-52.

226. Sasieni P, Castanon A, Cuzick J. Effectiveness of cervical screening with age: population based case-control study of prospectively recorded data. *BMJ* 2009;**339**: b2968.

227. Arbyn M, Van Oyen H. Cervical cancer screening in Belgium. *Eur J Cancer* 2000;**36**: 2191-7.

228. De Vuyst H, Clifford G, Li N, Franceschi S. HPV infection in Europe. *Eur J Cancer* 2009;**45**: 2632-9.

229. Arbyn M, Benoy IH, Simoens C, et al. Pre-vaccination distribution of HPV types in women attending at cervical cancer screening in Belgium. *Cancer Epidemiol Biomarkers Prev* 2009;**20**: 321-30.

230. Cuzick J, Clavel C, Petry KU, et al. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. *Int J Cancer* 2006;**119**: 1095-101.

231. Schneider A, Hoyer H, Lotz B, et al. Screening for high-grade cervical intra-epithelial neoplasia and cancer by testing for high-risk HPV, routine cytology or colposcopy. *Int J Cancer* 2000;**89**: 529-34.

232. Cuzick J, Beverley E, Ho L, et al. HPV testing in primary screening of older women. *Br J Cancer* 1999;**81**: 554-8.

233. Bulkmans NW, Rozendaal L, Snijders PJ, et al. POBASCAM, a population-based randomized controlled trial for implementation of high-risk HPV testing in cervical screening: design, methods and baseline data of 44,102 women. *Int J Cancer* 2004;**110**: 94-101.

234. Kulasingam SL, Hughes JP, Kiviat NB, et al. Evaluation of human papillomavirus testing in primary screening cervical abnormalities. Comparison of sesitivity, specificity, and frequency of referral. *JAMA* 2002;**288**: 1749-57.

235. Ratnam S, Franco EL, Ferenczy A. Human papillomavirus testing for primary screening of cervical cancer precursors. *Cancer Epidemiol Biomarkers Prev* 2000;**9**: 945-51.

236. Arbyn M, Ronco G, Cuzick J, Wentzensen N, Castle PE. How to evaluate emerging technologies in cervical cancer screening? *Int J Cancer* 2009;**125**: 2489-96.

237. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010;**11**: 249-57.

238. Ronco G, Biggeri A, Confortini M, et al. [Health Technology assessment Report: HPV Dna based primary screening for cervical cancer precursors]. *Epidemiol Prev* 2012;**36**: 1-72.

239. European Commission. European Guidelines for Quality Assurance in Cervical Cancer Screening: Supplements to the 2nd Edition.Editors: Anttila A, Arbyn M, De Vuyst H, Dillner J, Franceschi S, Patnick J et al. Luxembourg: Office for Official Publications of the European Communities; 2015.

240. Bruinsma F, Quinn M. The risk of preterm birth following treatment for precancerous changes in the cervix: a systematic review and meta-analysis. *BJOG* 2011;**118**: 1031-41.

241. Noehr B, Jensen A, Frederiksen K, Tabor A, Kjaer SK. Depth of cervical cone removed by loop electrosurgical excision procedure and subsequent risk of spontaneous preterm delivery. *Obstet Gynecol* 2009;**114**: 1232-8.

242. Castanon A, Brocklehurst P, Evans H, et al. Risk of preterm birth after treatment for cervical intraepithelial neoplasia among women attending colposcopy in England: retrospective-prospective cohort study. *BMJ* 2012;**345**: e5174.

243. Werner CL, Lo JY, Heffernan T, et al. Loop electrosurgical excision procedure and risk of preterm birth. *Obstet Gynecol* 2010;**115**: 605-8.

244. Kyrgiou M, Athanasiou A, Paraskevaidi M, et al. Adverse obstetric outcomes after local treatment for cervical preinvasive and early invasive disease according to cone depth: systematic review and meta-analysis. *BMJ* 2016;**354**: i3633.

245. Habbema D, Weinmann S, Arbyn M, et al. Harms of cervical cancer screening in the United States and the Netherlands. *Int J Cancer* 2016; in-press.

246. Arbyn M, de Sanjose S, Saraiya M, et al. EUROGIN 2011 roadmap on prevention and treatment of HPV-related disease. *Int J Cancer* 2012;**131**: 1969-82.

247. Meijer CJLM, Castle PE, Hesselink AT, 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**: 516-20.

248. Snijders PJ, van den Brule AJ, Meijer CJLM. The clinical relevance of human papillomavirus testing : relationship between analytical and clinical sensitivity. *J Pathol* 2003;**201**: 1-6.

249. Tang NS, Tang ML, Chan IS. On tests of equivalence via non-unity relative risk for matched-pair design. *Stat Med* 2003;**22**: 1217-33.

250. Arbyn M, Snijders PJ, Meijer CJLM, et al. Which high-risk HPV assays fulfil criteria for use in primary cervical cancer screening? *Clin Microbiol Infect* 2015;**21**: 817-26.

251. Arbyn M, Depuydt C, Benoy I, et al. VALGENT: a protocol for clinical validation of human papillomavirus assays. *J Clin Virol* 2016;**76 (Suppl 1)**: S14-S21.

252. Heideman DA, Hesselink AT, van Kemenade FJ, et al. The APTIMA HPV assay fulfills the cross-sectional clinical and reproducibility criteria of international guidelines for HPV test requirements for cervical screening. *J Clin Microbiol* 2013;**51**: 3653-7.

253. Carozzi FM, Burroni E, Bisanzi S, et al. Comparison of Clinical Performance of Abbott RealTime High Risk HPV Test with That of Hybrid Capture 2 Assay in a Screening Setting. *J Clin Microbiol* 2011;**49**: 1446-51.

254. Poljak M, Ostrbenk A, Seme K, et al. Comparison of Clinical and Analytical Performance of the Abbott RealTime High Risk HPV Test to the Performance of Hybrid Capture 2 in Population-Based Cervical Cancer Screening. *J Clin Microbiol* 2011;**49**: 1721-9.

255. Hesselink AT, Meijer CJLM, Poljak M, et al. Clinical validation of the Abbott RealTime High Risk (HR) HPV assay according to the guidelines for HPV DNA test requirements for cervical screening. *J Clin Microbiol* 2013;**51**: 2409-10.

256. Ejegod DM, Serrano I, Cuschieri K, et al. Clinical Validation of the BD Onclarity™ HPV Assay Using a Non-Inferiority Test. *J Med Microbiol Diagn* 2014;**S3: 003**: 1-4.

257. Boers A, Wang R, Slagter-Menkema L, et al. Clinical validation of the Cervista HPV HR test according to the international guidelines for human papillomavirus test requirements for cervical cancer screening. *J Clin Microbiol* 2014;**52**: 4391-3.

258. Alameda F, Garrote L, Mojal S, et al. Cervista HPV HR Test for Cervical Cancer Screening: A Comparative Study in the Catalonian Population. *Arch Pathol Lab Med* 2015;**139**: 241-4.

259. Heideman DA, Hesselink AT, Berkhof J, et al. Clinical validation of the cobas(R)4800 HPV Test for cervical screening purposes. *J Clin Microbiol* 2011;**49**: 3983-5.

260. Lloveras B, Gomez S, Alameda F, et al. HPV Testing by cobas HPV Test in a Population from Catalonia. *PLoS ONE* 2013;**8**: e58153.

261. Depuydt CE, Benoy IH, Beert JF, et al. Clinical validation of a type-specific real time quantitative human papillomavirus PCR to the performance of Hybrid Capture 2 for the purpose of cervical cancer screening. *J Clin Microbiol* 2012;**50**: 4073-7.

262. Hesselink A, Berkhof J, van der Salm ML, et al. Clinical validation of the HPV-Risk assay: a novel, real-time PCR assay for the detection of high-risk human papillomavirus DNA by targeting the E7 region. *J Clin Microbiol* 2014;**52**: 890-6.

263. Hesselink AT, Heideman DA, Berkhof J, et al. Comparison of the clinical performance of PapilloCheck human papillomavirus detection with that of the GP5+/6+-PCR-enzyme immunoassay in population-based cervical screening. *J Clin Microbiol* 2010;**48**: 797-801.

264. Kwon MJ, Roh KH, Park H, Woo HY. Comparison of the Anyplex II HPV28 assay with the Hybrid Capture 2 assay for the detection of HPV infection. *J Clin Virol* 2014;**59**: 246-9.

265. Geraets D, Cuschieri K, Koning M, et al. Clinical evaluation of a GP5+/6+-based Luminex assay having full high-risk HPV genotyping capability and an internal control. *J Clin Microbiol* 2014;**52**: 3996-4002.

266. Qiao YL, Sellors JW, Eder PS, et al. A new HPV-DNA test for cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. *Lancet Oncol* 2008;**9**: 929-36.

267. Belinson JL, Wu R, Belinson SE, et al. A population-based clinical trial comparing endocervical high-risk HPV testing using hybrid capture 2 and Cervista from the SHENCCAST II Study. *Am J Clin Pathol* 2011;**135**: 790-5.

268. Hovland S, Arbyn M, Lie AK, et al. A comprehensive evaluation of the accuracy of cervical pre-cancer detection methods in a high-risk area in East Congo. *Br J Cancer* 2010;**102**: 957-65.

269. Einstein MH, Smith KM, Davis TE, et al. Clinical Evaluation of the Cartridge-Based GeneXpert Human Papillomavirus Assay in Women Referred for Colposcopy. *J Clin Microbiol* 2014;**in press**: in-press.

270. Heard I, Cuschieri K, Geraets DT, Quint W, Arbyn M. Clinical and analytical performance of the PapilloCheck HPV-Screening assay using the VALGENT framework. *J Clin Virol* 2016;**81**: 6-11.

271. Cuschieri K, Geraets DT, Moore C, et al. Clinical and analytical performance of the Onclarity HPV assay using the VALGENT framework. *J Clin Microbiol* 2015;**53**: 3272-9.

272. Ejegod D, Bottari F, Pedersen H, Sandri MT, Bonde J. The BD Onclarity HPV assay on SurePath collected samples meets the International Guidelines for Human Papillomavirus Test Requirements for Cervical Screening. *J Clin Microbiol* 2016;**54**: 2267-72.

273. Hesselink AT, Sahli R, Berkhof J, et al. Clinical validation of Anyplex II HPV HR Detection according to the guidelines for HPV test requirements for cervical cancer screening. *J Clin Virol* 2016;**76**: 36-9.

274. Jung S, Lee B, Lee KN, Kim Y, Oh EJ. Clinical Validation of Anyplex II HPV HR Detection Test for Cervical Cancer Screening in Korea. *Arch Pathol Lab Med* 2016;**140**: 276-80.

275. Cuschieri K, Geraets D, Cuzick J, et al. Performance of a cartridge based assay for the detection of clinically significant HPV infection - lessons from VALGENT (Validation of HPV Genotyping Tests). *J Clin Microbiol* 2016;**54**: 2337-47.

276. Morrison EAB, Goldberg GL, Hagan RJ, Kadish AS, Burk RD. Self-administered home cervicovaginal lavage: A novel tool for the clinical-epidemiologic investigation of genital human Papillomavirus infection. *Am J Obstet Gynecol* 1992;**167**: 104-7.

277. Hillemanns P, Kimmig R, Hüttemann U., Dannecker C., Thaler C.J. Screening for cervical neoplasia by self-assessment for human papillomavirus. *Lancet* 1999;**354**: 1970.

278. Sellors JW, Lorincz AT, Mahony JB, et al. Comparison of self-collected vaginal, vulvar and urine samples with physician-collected cervical samples for human papillomavirus testing to detect high-grade sqaumous intraepethelial lesions. *CMAJ* 2000;**163**: 513-8.

279. Wright TC, Denny L, Kuhn L, Pollack A, Lorincz AT. HPV DNA testing of self-collected vaginal samples compared with cytologic screening to detect cervical cancer. *JAMA* 2000;**283**: 81-6.

280. Lorenzato FR, Singer A, Ho L, et al. Human papillomavirus detection for cervical cancer prevention with polymerase chain reaction in self-collected samples. *Am J Obstet Gynecol* 2002;**186**: 962-8.

281. Nobbenhuis MA, Helmerhorst TJ, van den Brule AJ, et al. Primary screening for high risk HPV by home obtained cervicovaginal lavage is an alternative screening tool for unscreened women. *J Clin Pathol* 2002;**55**: 435-9.

282. Garcia F, Barker B, Santos C, et al. Cross-sectional study of patient- and physician-collected cervical cytology and human papillomavirus. *Obstet Gynecol* 2003;**102**: 266-72.

283. Salmerón J, Lazcano Ponce EC, Lorincz AT, et al. Comparison of HPV-based assays with Papaniclaou smears for cervical cancer screening in Morelos State, Mexico. *Cancer Causes Control* 2003;**14**: 505-12.

284. Brink AA, Meijer CJLM, Wiegerinck MA, et al. High concordance of results of testing for human papillomavirus in cervicovaginal samples collected by two methods, with comparison of a novel self-sampling device to a conventional endocervical brush. *J Clin Microbiol* 2006;**44**: 2518-23.

285. Daponte A, Pournaras S, Mademtzis I, et al. Evaluation of HPV 16 PCR detection in self- compared with clinician-collected samples in women referred for colposcopy. *Gynecol Oncol* 2006;**103**: 463-6.

286. Girianelli VR, Thuler LC, Szklo M, et al. Comparison of human papillomavirus DNA tests, liquid-based cytology and conventional cytology for the early detection of cervix uteri cancer. *Eur J Cancer Prev* 2006;**15**: 504-10.

287. Holanda F, Jr., Castelo A, Veras TM, et al. Primary screening for cervical cancer through self sampling. *Int J Gynaecol Obstet* 2006;**95**: 179-84.

288. Seo SS, Song YS, Kim JW, et al. Good correlation of HPV DNA test between self-collected vaginal and clinician-collected cervical samples by the oligonucleotide microarray. *Gynecol Oncol* 2006;**102**: 67-73.

289. Szarewski A, Cadman L, Mallett S, et al. Human papillomavirus testing by self-sampling: assessment of accuracy in an unsupervised clinical setting. *J Med Screen* 2007;**14**: 34-42.

290. Bhatla N, Dar L, Patro AR, et al. Can human papillomavirus DNA testing of self-collected vaginal samples compare with physician-collected cervical samples and cytology for cervical cancer screening in developing countries? *Cancer Epidemiol* 2009;**33**: 446-50.

291. Balasubramanian A, Kulasingam SL, Baer A, et al. Accuracy and cost-effectiveness of cervical cancer screening by high-risk human papillomavirus

DNA testing of self-collected vaginal samples. *J Low Genit Tract Dis* 2010;**14**: 185-95.

292. Gustavsson I, Sanner K, Lindell M, et al. Type-specific detection of high-risk human papillomavirus (HPV) in self-sampled cervicovaginal cells applied to FTA elute cartridge. *J Clin Virol* 2011;**51**: 255-8.

293. Lazcano-Ponce E, Lorincz AT, Cruz-Valdez A, et al. Self-collection of vaginal specimens for human papillomavirus testing in cervical cancer prevention (MARCH): a community-based randomised controlled trial. *Lancet* 2011;**378**: 1868-73.

294. Taylor S, Wang C, Wright TC, Denny L, Kuhn L. A comparison of human papillomavirus testing of clinician-collected and self-collected samples during follow-up after screen-and-treat. *Int J Cancer* 2011;**129**: 879-86.

295. Twu NF, Yen MS, Lau HY, et al. Type-specific human papillomavirus DNA testing with the genotyping array: a comparison of cervical and vaginal sampling. *Eur J Obstet Gynecol Reprod Biol* 2011;**156**: 96-100.

296. Wikstrom I, Lindell M, Sanner K, Wilander E. Self-sampling and HPV testing or ordinary Pap-smear in women not regularly attending screening: a randomised study. *Br J Cancer* 2011;**105**: 337-9.

297. Belinson JL, Du H, Yang B, et al. Improved sensitivity of vaginal self-collection and high-risk human papillomavirus testing. *Int J Cancer* 2012;**130**: 1855-60.

298. Dijkstra MG, Heideman DA, van Kemenade FJ, et al. Brush-based self-sampling in combination with GP5+/6+-PCR-based hrHPV testing: high concordance with physician-taken cervical scrapes for HPV genotyping and detection of high-grade CIN. *J Clin Virol* 2012;**54**: 147-51.

299. Longatto-Filho A, Naud P, Derchain SF, et al. Performance characteristics of Pap test, VIA, VILI, HR-HPV testing, cervicography, and colposcopy in diagnosis of significant cervical pathology. *Virchows Arch* 2012;**460**: 577-85.

300. van Baars R, Bosgraaf RP, ter Harmsel BW, et al. Dry storage and transport of a cervicovaginal self-sample using the Evalyn Brush(R): reliable HPV detection combined with women's comfort. *J Clin Microbiol* 2012;**50**: 3937-43.

301. Zhao FH, Lewkowitz AK, Chen F, et al. Pooled Analysis of a Self-Sampling HPV DNA Test as a Cervical Cancer Primary Screening Method. *J Natl Cancer Inst* 2012;**104**: 178-88.

302. Darlin L, Borgfeldt C, Forslund O, et al. Vaginal self-sampling without preservative for human papillomavirus testing shows good sensitivity. *J Clin Virol* 2013;**56**: 52-6.

303. Geraets DT, van BR, Alonso I, et al. Clinical evaluation of high-risk HPV detection on self-samples using the indicating FTA-elute solid-carrier cartridge. *J Clin Virol* 2013;**57**: 125-9.

304. Guan Y, Gravitt PE, Howard R, et al. Agreement for HPV genotyping detection between self-collected specimens on a FTA cartridge and clinician-collected specimens. *J Virol Methods* 2013;**189**: 167-71.

305. Jentschke M, Soergel P, Hillemanns P. Evaluation of a multiplex real time PCR assay for the detection of human, papillomavirus infections on self-collected cervicovaginal lavage samples. *J Virol Methods* 2013;**193**: 131-4.

306. Jentschke M, Lange V, Soergel P, Hillemanns P. Enzyme-linked immunosorbent assay for p16(INK4a) - a new triage test for the detection of cervical intraepithelial neoplasia? *Acta Obstet Gynecol Scand* 2013;**92**: 160-4.

307. Nieves L, Enerson CL, Belinson S, et al. Primary cervical cancer screening and triage using an mRNA human papillomavirus assay and visual inspection. *Int J Gynecol Cancer* 2013;**23**: 513-8.

308. Arbyn M, Verdoodt F, Snijders PJF, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. *Lancet Oncol* 2014;**15**: 172-83.

309. Arbyn M, Castle P. 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;**24**: 769-72.

310. Belinson JL, Hu S, Niyazi M, et al. Prevalence of type-specific human papillomavirus in endocervical, upper and lower vaginal, perineal and vaginal self-collected specimens: Implications for vaginal self-collection. *Int J Cancer* 2010;**127**: 1151-7.

311. Cuzick J, Sasieni P, Davies P, et al. A systematic review of the role of human papillomavirus testing within a cervical screening programme. *Health Technol Assess* 1999;**3**: 1-204.

312. Wright TC, Jr., Massad LS, Dunton CJ, et al. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. *Am J Obstet Gynecol* 2007;**197**: 346-55.

313. Wentzensen N, Bergeron C, Cas F, Vinokurova S, von Knebel DM. Triage of women with ASCUS and LSIL cytology: use of qualitative assessment of p16INK4a positive cells to identify patients with high-grade cervical intraepithelial neoplasia. *Cancer* 2007;**111**: 58-66.

314. Wentzensen N, von Knebel DM. Biomarkers in cervical cancer screening. *Dis Markers* 2007;**23**: 315-30.

315. Schmidt D, Bergeron C, Denton KJ, Ridder R. p16/ki-67 dual-Stain cytology in the triage of ASCUS and LSIL papanicolaou cytology: Results from the european equivocal or mildly abnormal papanicolaou cytology study. *Cancer Cytopathol* 2011;**119**: 158-66.

316. Zuna RE, Wang SS, Rosenthal DL, et al. Determinants of human papillomavirus-negative, low-grade squamous intraepithelial lesions in the atypical squamous cells of undetermined significance/low-grade squamous intraepithelial lesions triage study (ALTS). *Cancer* 2005;**105**: 253-62.

317. Bais AG, Rebolj M, Snijders PJ, et al. Triage using HPV-testing in persistent borderline and mildly dyskaryotic smears: Proposal for new guidelines. *Int J Cancer* 2005;**116**: 122-9.

318. Berkhof J, de Bruijne MC, Zielinski GD, et al. Evaluation of cervical screening strategies with adjunct high-risk human papillomavirus testing for women with borderline or mild dyskaryosis. *Int J Cancer* 2006;**118**: 1759-68.

319. Xu L, Verdoodt F, Wentzensen N, Bergeron C, Arbyn M. Triage of ASC-H: A meta-analysis of the accuracy of high-risk HPV testing and other markers to detect cervical precancer. *Cancer Cytopathol* 2016;**124**: 261-72.

320. Soutter WP, Sasieni P, Panoskaltsis T. Long-term risk of invasive cervical cancer after treatment of squamous cervical intraepithelial neoplasia. *Int J Cancer* 2006;**118**: 2048-55.

321. Strander B, Ryd W, Wallin KL, et al. Does HPV-status 6-12 months after treatment of high grade dysplasia in the uterine cervix predict long term recurrence? *Eur J Cancer* 2007;**43**: 1849-55.

322. Kocken M, Helmerhorst TJ, Berkhof J, et al. Risk of recurrent high-grade cervical intraepithelial neoplasia after successful treatment: a long-term multi-cohort study. *Lancet Oncol* 2011;**12**: 441-50.

323. Cuzick J, Arbyn M, Ronco G, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. *Vaccine* 2008;**26 Suppl 10**: K29-K41.

324. Saslow D, Solomon D, Lawson HW, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin* 2012.

325. Castle PE, Rodriguez AC, Burk RD, et al. Short term persistence of human papillomavirus and risk of cervical precancer and cancer: population based cohort study. *BMJ* 2009;**339**: b2569-b2584.

326. Khan MJ, Castle PE, Lorincz AT, et al. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. *J Natl Cancer Inst* 2005;**97**: 1072-9.

327. Verdoodt F, Szarewski A, Halfon P, Cuschieri K, Arbyn M. Triage of women with minor abnormal cervical cytology: meta-analysis of the accuracy of an assay targeting mRNA of five high-risk HPV types. *Cancer Cytopathol* 2013; in-press.

328. Monsonego J, Hudgens MG, Zerat L, et al. Evaluation of oncogenic human papillomavirus RNA and DNA tests with liquid based cytology in primary cervical cancer screening (The FASE study). *Int J Cancer* 2011;**129**: 691-701.

329. Rebolj M, Bonde J, Njor SH, Lynge E. Human papillomavirus testing in primary cervical screening and the cut-off level for hybrid capture 2 tests: systematic review. *BMJ* 2011;**342**: d2757.

330. Blatt AJ, Kennedy R, Luff RD, Austin RM, Rabin DS. Comparison of cervical cancer screening results among 256,648 women in multiple clinical practices. *Cancer Cytopathol* 2015;**123**: 282-8.

331. Tao X, Griffith CC, Zhou X, et al. History of high-risk HPV and Pap test results in a large cohort of patients with invasive cervical carcinoma: Experience from the largest women's hospital in China. *Cancer Cytopathol* 2015.

332. Zheng B, Li Z, Griffith CC, et al. Prior high-risk HPV testing and Pap test results for 427 invasive cervical cancers in China's largest CAP-certified laboratory. *Cancer Cytopathol* 2015;**123**: 428-34.

333. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998;**17**: 2815-34.

334. Dijkstra MG, van ZM, Rozendaal L, et al. Safety of extending screening intervals beyond five years in cervical screening programmes with testing for high risk human papillomavirus: 14 year follow-up of population based randomised cohort in the Netherlands. *BMJ* 2016;**355**: i4924.

335. Bruni L, Diaz M, Barrionuevo-Rosas L, et al. Global estimates of human papillomavirus vaccination coverage by region and income level: a pooled analysis. *Lancet Glob Health* 2016;**4**: e453-e463.

336. Drolet M, Benard E, Boily MC, et al. Population-level impact and herd effects following human papillomavirus vaccination programmes: a systematic review and meta-analysis. *Lancet Infect Dis* 2015;**15**: 565-80.

337. Palmer TJ, McFadden M, Pollock KG, et al. HPV immunisation and cervical screening-confirmation of changed performance of cytology as a screening test in immunised women: a retrospective population-based cohort study. *Br J Cancer* 2016;**114**: 582-9.

338. Franco EL, Cuzick J. Cervical cancer screening following prophylactic human papillomavirus vaccination. *Vaccine* 2008;**26 Suppl 1**: A16-A23.

339. Giorgi-Rossi P, Carozzi F, Federici A, et al. Cervical cancer screening in women vaccinated against human papillomavirus infection: Recommendations from a consensus conference. *Prev Med* 2017; in-press.

340. Canfell K, Chesson H, Kulasingam SL, et al. Modeling Preventative Strategies against Human Papillomavirus-Related Disease in Developed Countries. *Vaccine* 2012;**30 Suppl 5**: F157-F167.

341. de Kok IM, van Rosmalen J, Dvan Rosmalen J, et al. Primary screening for human papillomavirus compared with cytology screening for cervical cancer in European settings: cost effectiveness analysis based on a Dutch microsimulation model. *BMJ* 2012;**344**: e670.

342. Kim JJ, Burger EA, Sy S, Campos NG. Optimal Cervical Cancer Screening in Women Vaccinated Against Human Papillomavirus. *J Natl Cancer Inst* 2017;**109**.

343. Lew JB, Simms KT, Smith MA, et al. Primary HPV testing versus cytology-based cervical screening in women in Australia vaccinated for HPV and unvaccinated: effectiveness and economic assessment for the National Cervical Screening Program. *Lancet Pub Health* 2017;**2**: e96-e107.

344. Lew JB, Simms K, Smith M, et al. Effectiveness Modelling and Economic Evaluation of Primary HPV Screening for Cervical Cancer Prevention in New Zealand. *PLoS ONE* 2016;**11**: e0151619.

345. 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;**104**: e159-e166.

346. Jentschke M, Chen K, Arbyn M, et al. Direct comparison of two vaginal self-sampling devices for the detection of human papillomavirus infections. *J Clin Virol* 2016;**82**: 46-50.

347. Snellenberg S, De Strooper LM, Hesselink AT, et al. Development of a multiplex methylation-specific PCR as candidate triage test for women with an HPV-positive cervical scrape. *BMC Cancer* 2012;**12**: 551-*.

348. Wentzensen N, Sun C, Ghosh A, et al. Methylation of HPV18, HPV31, and HPV45 Genomes and Cervical Intraepithelial Neoplasia Grade 3. *J Natl Cancer Inst* 2012.

349. Schunemann HJ, Oxman AD, Brozek J, et al. GRADE: Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;**336**: 1106-10.

350. Nyaga VN, Arbyn M, Aerts M. Beta-binomial ANOVA model for network meta-analysis of diagnostic test accuracy data. *Stat Methods Med Res* 2016; accepeted for-publication.

351. Nyaga VN, Arbyn M, Aerts M. CopulaDTA: Copula based bivariate beta-binomial models for diagnostic test accuracy studies in a Bayesian framework. *J Stat Softw* 2016.

352. Nyaga VN, Aerts M, Arbyn M. ANOVA model for network meta-analysis of diagnostic test accuracy data. *Stat Methods Med Res* 2016; pre-pub.

353. Riley RD, Ahmed I, Debray TP, et al. Summarising and validating test accuracy results across multiple studies for use in clinical practice. *Stat Med* 2015;**34**: 2081-103.

354. Arbyn M, Redman CWE, Verdoodt F, et al. Incomplete excision of cervical pre-cancer as predictor of treatment failure: a systematic review and meta-analysis. *Lancet Oncol* 2017; in-press.

355. Cornet K. Kan de dekkingsgraad voor screening van baarmoederhalskanker verhoogd worden door gebruik te maken van een thuistest voor screening: een pilootstudie in een huisartsenpraktijk? Vrije Universiteit Brussel; 2015.

356. Arbyn M, Cuzick J. International agreement to join forces in synthesizing evidence on new methods for cervical cancer prevention. *Cancer Lett* 2009;**278**: 1-2.

357. Herbert A, Bergeron C, Wiener H, et al. European guidelines for quality assurance in cervical cancer screening: recommendations for cervical cytology terminology. *Cytopathology* 2007;**18**: 213-9.

358. Luff RD, National Cancer Institute. The revised Bethesda system for reporting cervical/vaginal cytological diagnoses. Report of the 1991 Bethesda Workshop. *Acta Cytol* 1992;**36**: 273-6.

359. Darragh TM, Colgan TJ, Cox JT, et al. The Lower Anogenital Squamous Terminology Standardization Project for HPV-Associated Lesions: Background and Consensus Recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. *Arch Pathol Lab Med* 2012;**136**: 1266-197.

360. Lorincz AT. Methods of DNA hybridisation and their clinical applicability to human papillomavirus detection. In: Franco EL, Monsonego J, editors. New developments in cervican cancer screening and prevention. Blackwell Science ed. Oxford: 1997. p. 325-37.

361. Jacobs MV, Snijders PJ, van den Brule AJ, et al. A general primer GP5+/GP6+ -mediated PCR-enzyme immunoassay method for rapid detection of 14 high risk and 6 low risk human papillomavirus types in cervical scrapings. *J Clin Microbiol* 1997;**35**: 791-5.

362. Cornall AM, Poljak M, Garland SM, et al. Anyplex II HPV28 detection and Anyplex II HPV HR detection assays are highly concordant with other commercial assays for detection of high-risk HPV genotypes in women with high grade cervical abnormalities. *Eur J Clin Microbiol Infect Dis* 2016.

363. Latsuzbaia A, Tapp J, Nguyen T, et al. Analytical performance evaluation of Anyplex II HPV28 and Euroarray HPV for genotyping of cervical samples. *Diagn Microbiol Infect Dis* 2016;**85**: 318-22.

364. Dockter J, Schroder A, Eaton B, et al. Analytical characterization of the APTIMA HPV Assay. *J Clin Virol* 2009;**45 Suppl 1**: S39-S47.

365. de Arruda M, Lyamichev VI, Eis PS, et al. Invader technology for DNA and RNA analysis: principles and applications. *Expert Rev Mol Diagn* 2002;**2**: 487-96.

366. Cornall AM, Poljak M, Garland SM, et al. EUROarray human papillomavirus (HPV) assay is highly concordant with other commercial assays for detection of high-risk HPV genotypes in women with high grade cervical abnormalities. *Eur J Clin Microbiol Infect Dis* 2016;**35**: 1033-6.

367. Micalessi MI, Boulet GA, Vorsters A, et al. A real-time PCR approach based on SPF10 primers and the INNO-LiPA HPV Genotyping Extra assay for the detection and typing of human papillomavirus. *J Virol Methods* 2013;**187**: 166-71.

368. Castle PE, Gravitt PE, Solomon D, Wheeler CM, Schiffman M. Comparison of linear array and line blot assay for detection of human papillomavirus and diagnosis of cervical precancer and cancer in the atypical squamous cell of undetermined significance and low-grade squamous intraepithelial lesion triage study. *J Clin Microbiol* 2008;**46**: 109-17.

369. Soderlund-Strand A, Dillner J, Carlson J. High-throughput genotyping of oncogenic human papilloma viruses with maldi-tof mass spectrometry. *Clin Chem* 2008;**54**: 86-92.

370. Castle PE, Gutierrez EC, Leitch SV, et al. An Evaluation of a New DNA Test for Carcinogenic Human Papillomavirus. *J Clin Microbiol* 2011;**49**: 3029-32.

371. Molden T, Kraus I, Skomedal H, Nordstrom T, Karlsen F. PreTect HPV-Proofer: real-time detection and typing of E6/E7 mRNA from carcinogenic human papillomaviruses. *J Virol Methods* 2007;**142**: 204-12.

372. Castle PE, Smith KM, Davis TE, et al. Reliability of the Xpert HPV assay to detect high-risk human papillomavirus DNA in a colposcopy referral population. *Am J Clin Pathol* 2015;**143**: 126-33.

373. Manos MM, Kinney WK, Hurley LB, et al. Identifying Women With Cervical Neoplasia: Using Human Papillomavirus DNA Testing for Equivocal Papanicolaou Results. *JAMA* 1999;**281**: 1605-10.

374. Bergeron C, Jeannel D, Poveda J, Cassonnet P, Orth G. Human papillomavirus testing in women with mild cytologic atypia. *Obstet Gynecol* 2000;**95**: 821-7.

375. Lytwyn A, Sellors JW, Mahony JB. Comparison of human papillomavirus DNA testing and repeat Papanicolaou test in women with low-grade cervical cytologic abnormalities: a randomized trial. *CMAJ* 2000;**163**: 701-7.

376. Shlay JC, Dunn T, Byers T, Baron AE, Douglas JM, Jr. Prediction of cerviacal intraepithelial neoplasia grade 2-3 using risk assessement and human papillomavirus testing in women with atypia on papanicolaou smears. *Obstet Gynecol* 2000;**96**: 410-6.

377. Morin C, Bariati C, Bouchard C, et al. Managing atypical squamous cells of undetermined significance in Papanicolaou smears. *J Reprod Med* 2001;**46**: 799-805.

378. Rebello G, Hallam N, Smart G, Farquharson D, McCafferty J. Human papillomavirus testing and the management of women with mildly abnormal cervical smears: an observational study. *BMJ* 2001;**322**: 894-5.

379. Solomon D, Schiffman MA, Tarone B. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance (ASCUS): baseline results from a randomized trial. *J Natl Cancer Inst* 2001;**93**: 293-9.

380. Zielinski GD, Snijders PJ, Rozendaal L, et al. High-risk HPV testing in women with borderline and mild dyskaryosis: long-term follow-up data and clinical relevance. *J Pathol* 2001;**195**: 300-6.

381. Pretorius RG, Peterson P, Novak S, et al. Comparison of two signal-amplification DNA tests for high-risk HPV as an aid to colposcopy. *J Reprod Med* 2002;**47**: 290-6.

382. Cuzick J, Szarewski A, Cubie H, et al. Management of women who test positive for high-risk types of human papillomavirus: the HART study. *Lancet* 2003;**362**: 1871-6.

383. Guyot A, Karim S, Kyi MS, Fox J. Evaluation of adjunctive HPV testing by hybrid capture II(R) in women with minor cytological abnormalities for the diagnosis of CIN2/3 and cost comparison with colposcopy. *BMC Infect Dis* 2003;**3**: 1-7.

384. Lonky NM, Felix JC, Naidu YM, Wolde Tsadik G. Triage of atypical squamous cells of undetermined significance with Hybrid Capture II: colposcopy and histologic human papillomavirus correlation. *Obstet Gynecol* 2003;**101**: 481-9.

385. Ordi J, Puig-Tintore LM, Torne A, et al. [Contribution of high risk human papillomavirus testing to the management of premalignant and malignant lesions of the uterine cervix]. *Med Clin (Barc)* 2003;**121**: 441-5.

386. Wensveen C, Kagie M, Veldhuizen R, et al. Detection of cervical intraepithelial neoplasia in women with atypical squamous or glandular cells of undetermined significance cytology: a prospective study. *Acta Obstet Gynecol Scand* 2003;**82**: 883-9.

387. Andersson S, Dillner L, Elfgren K, et al. A comparison of the human papillomavirus test and Papanicolaou smear as a second screening method for women with minor cytological abnormalities. *Acta Obstet Gynecol Scand* 2005;**84**: 996-1000.

388. Chen HS, Yang YC, Su TH, Wang TY, Huang YW. Human papillomavirus testing (Hybrid Capture II) to detect high-grade cervical intraepithelial neoplasia in women with mildly abnormal Papanicolaou results. *Taiwan J Obstet Gynecol* 2005;**44**: 252-7.

389. Dalla Palma P, Pojer A, Girlando S. HPV triage of women with atypical squamous cells of undetermined significance: a 3 year experience in an Italian organized programme. *Cytopathology* 2005;**16**: 22-6.

390. Davis-Devine S, Day SJ, Freund GG. Test performance comparison of inform HPV and hybrid capture 2 high-risk HPV DNA tests using the SurePath liquid-based Pap test as the collection method. *Am J Clin Pathol* 2005;**124**: 24-30.

391. Giovannelli L, Capra G, Lama A, et al. Atypical squamous cells of undetermined significance-favour reactive compared to atypical squamous cells of undetermined significance-favour dysplasia: association with cervical intraepithelial lesions and human papillomavirus infection. *J Clin Virol* 2005;**33**: 281-6.

392. Nieh S, Chen SF, Chu TY, et al. Is p16(INK4A) expression more useful than human papillomavirus test to determine the outcome of atypical squamous cells of undetermined significance-categorized Pap smear? A comparative analysis using abnormal cervical smears with follow-up biopsies. *Gynecol Oncol* 2005;**97**: 35-40.

393. Bergeron C, Cas F, Fagnani F, et al. Assessment of human papillomavirus testing on liquid-based CYTO-screen system for women with atypical squamous cells of undetermined significance. Effect of age. *Gynecol Obstet Fertil* 2006;**34**: 312-6.

394. Holladay EB, Logan S, Arnold J, Knesel B, Smith GD. A comparison of the clinical utility of p16(INK4a) immunolocalization with the presence of human papillomavirus by hybrid capture 2 for the detection of cervical dysplasia/neoplasia. *Cancer* 2006;**108**: 451-61.

395. Kelly D, Kincaid E, Fansler Z, Rosenthal DL, Clark DP. Detection of cervical high-grade squamous intraepithelial lesions from cytologic samples using a novel immunocytochemical assay (ProExtrade mark C). *Cancer* 2006;**108**: 494-500.

396. Kiatpongsan S, Niruthisard S, Mutirangura A, et al. Role of human papillomavirus DNA testing in management of women with atypical squamous cells of undetermined significance. *Int J Gynecol Cancer* 2006;**16**: 262-5.

397. Monsonego J, Pintos J, Semaille C, et al. Human papillomavirus testing improves the accuracy of colposcopy in detection of cervical intraepithelial neoplasia. *Int J Gynecol Cancer* 2006;**16**: 591-8.

398. Cuschieri KS, Graham C, Moore C, Cubie HA. Human Papillomavirus testing for the management of low-grade cervical abnormalities in the UK-Influence of age and testing strategy. *J Clin Virol* 2007;**38**: 14-8.

399. Ronco G, Cuzick J, Segnan N, et al. HPV triage for low grade (L-SIL) cytology is appropriate for women over 35 in mass cervical cancer screening using liquid based cytology. *Eur J Cancer* 2007;**43**: 476-80.

400. You K, Liang X, Qin F, Guo Y, Geng L. High-risk human papillomavirus DNA testing and high-grade cervical intraepithelial lesions. *Aust N Z J Obstet Gynaecol* 2007;**47**: 141-4.

401. De Francesco MA, Gargiulo F, Schreiber C, et al. Comparison of the AMPLICOR Human Papillomavirus Test and the Hybrid Capture 2 Assay for detection of high-risk human papillomavirus in women with abnormal PAP smear. *J Virol Methods* 2008;**147**: 10-7.

402. Monsonego J, Pollini G, Evrard MJ, et al. Detection of Human Papillomavirus Genotypes Among High-Risk Women: A Comparison of Hybrid Capture and Linear Array Tests. *Sex Transm Dis* 2008;**35**: 521-7.

403. Siddiqui MT, Hornaman K, Cohen C, Nassar A. ProEx C immunocytochemistry and high-risk human papillomavirus DNA testing in papanicolaou tests with atypical squamous cell (ASC-US) cytology: correlation study with histologic biopsy. *Arch Pathol Lab Med* 2008;**132**: 1648-52.

404. Szarewski A, Ambroisine L, Cadman L, et al. Comparison of Predictors for High-Grade Cervical Intraepithelial Neoplasia in Women with Abnormal Smears. *Cancer Epidemiol Biomarkers Prev* 2008;**17**: 3033-43.

405. Cattani P, Zannoni GF, Ricci C, et al. Clinical performance of human papillomavirus E6 and E7 mRNA testing for high-grade lesions of the cervix. *J Clin Microbiol* 2009;**47**: 3895-901.

406. Huang S, Erickson B, Tang N, et al. Clinical performance of Abbott RealTime High Risk HPV test for detection of high-grade cervical intraepithelial neoplasia in women with abnormal cytology. *J Clin Virol* 2009;**45 Suppl 1**: S19-S23.

407. Lee JK, Kim MK, Song SH, et al. Comparison of human papillomavirus detection and typing by hybrid capture 2, linear array, DNA chip, and cycle sequencing in cervical swab samples. *Int J Gynecol Cancer* 2009;**19**: 266-72.

408. Silverloo I, Andrae B, Wilander E. Value of high-risk HPV-DNA testing in the triage of ASCUS. *Acta Obstet Gynecol Scand* 2009;**88**: 1006-10.

409. Del Mistro A, Frayle-Salamanca H, Trevisan R, et al. Triage of women with atypical squamous cells of undetermined significance (ASC-US): Results of an Italian multicentric study. *Gynecol Oncol* 2010;**117**: 77-81.

410. Halfon P, Benmoura D, Agostini A, et al. Stepwise algorithm combining HPV high-risk DNA-based assays and RNA-based assay for high grade CIN in women with abnormal smears reffered to colposcopy. *Cancer Biomarkers* 2010;**7**: 133-9.

411. Alameda F, Piujan L, Lioveras B, et al. The value of p16 in ASCUS cases: A retrospective study using frozen cytologic material. *Diagn Cytopathol* 2011;**39**: 110-4.

412. Clad A, Reuschenbach M, Weinschenk J, et al. Performance of the APTIMA high-risk HPV mRNA assay in a referral population in comparison with Hybrid Capture 2 and cytology. *J Clin Microbiol* 2011;**49**: 1071-6.

413. Dufresne S, Sauthier P, Mayrand MH, et al. Human papillomavirus (HPV) DNA triage of women with atypical squamous cells of undetermined significance with Amplicor HPV and Hybrid Capture 2 assays for detection of high-grade lesions of the uterine cervix. *J Clin Microbiol* 2011;**49**: 48-53.

414. Ratnam S, Coutlee F, Fontaine D, et al. Aptima HPV E6/E7 mRNA test is as sensitive as Hybrid Capture 2 Assay but more specific at detecting cervical precancer and cancer. *J Clin Microbiol* 2011;**49**: 557-64.

415. Stoler MH, Wright TC, Jr., Sharma A, et al. High-Risk Human Papillomavirus Testing in Women With ASC-US Cytology: Results From the ATHENA HPV Study. *Am J Clin Pathol* 2011;**135**: 468-75.

416. Szarewski A, Mesher D, Cadman L, et al. A comparison of seven tests for high grade cervical intraepithelial neoplasia in women with abnormal smears: the Predictors 2 study. *J Clin Microbiol* 2012;**50**: 1867-73.

417. Alaghehbandan R, Fontaine D, Bentley J, et al. Performance of ProEx C and PreTect HPV-proofer E6/E7 mRNA tests in comparison with the hybrid capture 2 HPV DNA test for triaging ASCUS and LSIL cytology. *Diagn Cytopathol* 2013;**41**: 767-75.

418. Oliveira A, Verdasca N, Pista A. Use of the NucliSENS EasyQHPV assay in the management of cervical intraepithelial neoplasia. *J Med Virol* 2013;**85**: 1235-41.

419. Monsonego J, Bohbot JM, Pollini G, et al. Performance of the Roche AMPLICOR(R) Human papillomavirus (HPV) test in prediction of cervical intraepithelial neoplasia (CIN) in women with abnormal PAP smear. *Gynecol Oncol* 2005;**99**: 160-8.

420. Wentzensen N, Gravitt PE, Solomon D, Wheeler CM, Castle PE. A study of Amplicor human papillomavirus DNA detection in the atypical squamous cells of undetermined significance-low-grade squamous intraepithelial lesion triage study. *Cancer Epidemiol Biomarkers Prev* 2009;**18**: 1341-9.

421. Jakobsson M, Tarkkanen J, Auvinen E, et al. Colposcopy referral rate can be reduced by high-risk human papillomavirus triage in the management of recurrent atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion cytology in Finland. *Int J STD AIDS* 2012;**23**: 485-9.

422. Halfon P, Benmoura D, Agostini A, et al. Evaluation of the clinical performance of the Abbott RealTime High-Risk HPV for carcinogenic HPV detection. *J Clin Virol* 2010;**48**: 246-50.

423. Froberg M, Johansson B, Hjerpe A, Andersson S. Human papillomavirus "reflex" testing as a screening method in cases of minor cytological abnormalities. *Br J Cancer* 2008;**99**: 563-8.

424. Halfon P, Benmoura D, Khiri H, et al. Comparison of the clinical performance of carcinogenic HPV typing of the Linear Array and Papillocheck HPV-screening assay. *J Clin Virol* 2010;**47**: 38-42.

425. Lapierre SG, Sauthier P, Mayrand MH, et al. Human Papillomavirus (HPV) DNA Triage of Women with Atypical Squamous Cells of Undetermined Significance with cobas 4800 HPV and Hybrid Capture 2 Tests for Detection of High-Grade Lesions of the Uterine Cervix. *J Clin Microbiol* 2012;**50**: 1240-4.

426. Dona MG, Vocaturo A, Giuliani M, et al. p16/Ki-67 dual staining in cervico-vaginal cytology: Correlation with histology, Human Papillomavirus detection

and genotyping in women undergoing colposcopy. *Gynecol Oncol* 2012;**126**: 198-202.

427. Waldstrom M, Ornskov D. Comparison of the clinical performance of an HPV mRNA test and an HPV DNA test in triage of atypical squamous cells of undetermined significance (ASC-US). *Cytopathology* 2012;**23**: 389-95.

428. Wentzensen N, Schwartz L, Zuna RE, et al. Performance of p16/Ki-67 immunostaining to detect cervical cancer precursors in a colposcopy referral population. *Clin Cancer Res* 2012;**18**: 4154-62.

429. Wong AK, Chan RC, Nichols WS, Bose S. Invader human papillomavirus (HPV) type 16 and 18 assays as adjuncts to HPV screening of cervical papanicolaou smears with atypical squamous cells of undetermined significance. *Cancer* 2009;**115**: 823-32.

430. Einstein MH, Martens MG, Garcia FAR, et al. Clinical validation of the Cervista« HPV HR and 16/18 genotyping tests for use in women with ASC-US cytology. *Gynecol Oncol* 2010;**118**: 116-22.

431. Molden T, Nygard JF, Kraus I, et al. Predicting CIN2+ when detecting HPV mRNA and DNA by PreTect HPV-proofer and consensus PCR: A 2-year follow-up of women with ASCUS or LSIL pap smear. *Int J Cancer* 2005;**114**: 973-6.

432. Andersson S, Hansson B, Norman I, et al. Expression of E6/E7 mRNA from 'high risk' human papillomavirus in relation to CIN grade, viral load and p16INK4a. *Int J Oncol* 2006;**29**: 705-11.

433. Halfon P, Benmoura D, Agostini A, et al. Relevance of HPV mRNA detection in a population of ASCUS plus women using the NucliSENS EasyQ HPV assay. *J Clin Virol* 2010;**47**: 177-81.

434. Sorbye SW, Fismen S, Gutteberg TJ, Mortensen ES. HPV mRNA test in women with minor cervical lesions; experience of the University Hospital of North Norway. *J Virol Methods* 2010;**169**: 219-22.

435. Andersson E, Karrberg C, Radberg T, et al. Type-specific HPV E6/E7 mRNA detection by real-time PCR improves identification of cervical neoplasia. *J Clin Microbiol* 2011;**49**: 3794-9.

436. Koliopoulos G, Chrelias C, Pappas A, et al. The diagnostic accuracy of two methods for E6&7 mRNA detection in women with minor cytological abnormalities. *Acta Obstet Gynecol Scand* 2012;**91**: 794-801.

437. Dockter J, Schroder A, Hill C, et al. Clinical performance of the APTIMA HPV Assay for the detection of high-risk HPV and high-grade cervical lesions. *J Clin Virol* 2009;**45**: S55-S61.

438. Wu R, Belinson SE, Du H, et al. Human papillomavirus messenger RNA assay for cervical cancer screening: the Shenzhen Cervical Cancer Screening Trial I. *Int J Gynecol Cancer* 2010;**20**: 1411-4.

439. Nieh S, Chen S-F, Chu T-Y, Lai H-C, Fu E. Expression of p16INK4A in Papanicolaou smears containing atypical squamous cells of undetermined significance from the uterine cervix. *Gynecol Oncol* 2003;**91**: 201-8.

440. Meyer JL, Hanlon DW, Andersen BT, Rasmussen OF, Bisgaard K. Evaluation of p16(INK4a) expression in ThinPrep cervical specimens with the CINtec p16(INK4a) assay: correlation with biopsy follow-up results. *Cancer* 2007;**111**: 83-92.

441. Monsonego J, Pollini G, Evrard MJ, et al. P16(INK4a) immunocytochemistry in liquid-based cytology samples in equivocal Pap smears: added value in management of women with equivocal Pap smear. *Acta Cytol* 2007;**51**: 755-66.

442. Schledermann D, Andersen BT, Bisgaard K, et al. Are adjunctive markers useful in routine cervical cancer screening? Application of p16(INK4a) and HPV-PCR on ThinPrep samples with histological follow-up. *Diagn Cytopathol* 2008;**36**: 453-9.

443. Guo M, Warriage I, Mutyala B, et al. Evaluation of p16 immunostaining to predict high-grade cervical intraepithelial neoplasia in women with Pap results of atypical squamous cells of undetermined significance. *Diagn Cytopathol* 2011;**39**: 482-8.

444. Passamonti B, Gustinuci D, Rechia P, et al. Expression of p16 in abnormal pap-tests as an indicator of CIN2+ lesions: a possible role in the low grade ASC/US and L/Sil (lg) cytologic lesions for screening prevention of uterine cervical tumours. *Pathologica* 2010;**102**: 6-11.

445. Samarawardana P, Dehn DL, Singh M, et al. p16(INK4a) is superior to high-risk human papillomavirus testing in cervical cytology for the prediction of underlying high-grade dysplasia. *Cancer Cytopathol* 2010;**118**: 146-56.

446. Sung CO, Kim SR, Oh YL, Song SY. The use of p16(INK4A) immunocytochemistry in "Atypical squamous cells which cannot exclude HSIL" compared with "Atypical squamous cells of undetermined significance" in liquid-based cervical smears. *Diagn Cytopathol* 2010;**38**: 168-71.

447. Nasioutziki M, Daniilidis A, Dinas K, et al. The evaluation of p16INK4a immunoexpression/immunostaining and human papillomavirus DNA test in cervical liquid-based cytological samples. *Int J Gynecol Cancer* 2011;**21**: 79-85.

448. Loghavi S, Walts AE, Bose S. CINtec(R) plus dual immunostain: A triage tool for cervical pap smears with atypical squamous cells of undetermined

significance and low grade squamous intraepithelial lesion. *Diagn Cytopathol* 2013;**41**: 582-7.

449. Edgerton N, Cohen C, Siddiqui MT. Evaluation of CINtec PLUS(R) testing as an adjunctive test in ASC-US diagnosed SurePath(R) preparations. *Diagn Cytopathol* 2013;**41**: 35-40.

450. Monsonego J, Pollini G, Evrard MJ, et al. Linear array genotyping and hybrid capture II assay in detecting human papillomavirus genotypes in women referred for colposcopy due to abnormal Papanicolaou smear. *Int J STD AIDS* 2008;**19**: 385-92.

451. Gage JC, Schiffman M, Solomon D, et al. Risk of precancer determined by HPV genotype combinations in women with minor cytologic abnormalities. *Cancer Epidemiol Biomarkers Prev* 2013; 1095-101.

452. Guo M, Lin CY, Gong Y, et al. Human papillomavirus genotyping for the eight oncogenic types can improve specificity of HPV testing in women with mildly abnormal Pap results. *Mod Pathol* 2008;**21**: 1037-43.

453. Depuydt CE, Makar AP, Ruymbeke MJ, et al. BD-ProExC as Adjunct Molecular Marker for Improved Detection of CIN2+ after HPV Primary Screening. *Cancer Epidemiol Biomarkers Prev* 2011;**20**: 628-37.

454. Spathis A, Kottaridi C, Chranioti A, et al. mRNA and DNA detection of human papillomaviruses in women of all ages attending two colposcopy clinics. *PLoS ONE* 2012;**7**: e49205-0.

455. Tambouret RH. [Diagnostic importance of ProEx C in gynecologic cytopathology]. *Ann Pathol* 2008;**28 Spec No 1**: S92-S93.

456. Lee NW, Kim D, Park JT, Kim A. Is the human papillomavirus test in combination with the Papanicolaou test useful for management of patients with diagnoses of atypical squamous cells of undetermined significance/low-grade squamous intraepithelial lesions? *Arch Pathol Lab Med* 2001;**125**: 1453-7.

457. Castle PE, Fetterman B, Thomas CJ, et al. The age-specific relationships of abnormal cytology and human papillomavirus DNA results to the risk of cervical precancer and cancer. *Obstet Gynecol* 2010;**116**: 76-84.

458. Voss JS, Kipp BR, Campion MB, et al. Assessment of fluorescence in situ hybridization and hybrid capture 2 analyses of cervical cytology specimens diagnosed as low grade squamous intraepithelial lesion for the detection of high grade cervical intraepithelial neoplasia. *Anal Quant Cytol Histol* 2010;**32**: 121-30.

459. Heider A, Austin RM, Zhao C. HPV Test Results Stratify Risk for Histopathologic Follow-Up Findings of High-Grade Cervical Intra-Epithelial

Neoplasia in Women with Low-Grade Squamous Intra-Epithelial Lesion Pap Results. *Acta Cytol* 2011;**55**: 48-53.

460. Levi AW, Harigopal M, Hui P, Schofield K, Chhieng DC. Use of high-risk human papillomavirus testing in patients with low-grade squamous intraepithelial lesions. *Cancer Cytopathol* 2011;**119**: 228-34.

461. Tsoumpou I, Valasoulis G, Founta C, et al. High-risk human papillomavirus DNA test and p16(INK4a) in the triage of LSIL: A prospective diagnostic study. *Gynecol Oncol* 2011;**121**: 49-53.

462. Ziemke P, Marquardt K. [Immunocytochemistry of p16(INK4a) and Ki-67 as adjunctive method for routine gynecological cytology of mild and moderate dysplasia.]. *Pathologe* 2013;**34**: 323-8.

463. Jones J, Powell NG, Tristram A, Fiander AN, Hibbitts S. Comparison of the PapilloCheck((R)) DNA micro-array Human Papillomavirus detection assay with Hybrid Capture II and PCR-enzyme immunoassay using the GP5/6+ primer set. *J Clin Virol* 2009;**45**: 100-4.

464. Cuzick J, Thomas CJ, Zhang G, et al. Human papillomavirus testing for triage of women with low-grade squamous intraepithelial lesions. *Int J Cancer* 2013;**132**: 959-66.