

VU Research Portal

Novel biomarkers for cervical screening and surveillance of women treated for cervical precancer

Uijterwaal, M.H.

2017

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Uijterwaal, M. H. (2017). *Novel biomarkers for cervical screening and surveillance of women treated for cervical precancer*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Chapter 5

High-risk human papillomavirus testing versus cytology in predicting post-treatment disease in women treated for high-grade cervical disease: A systematic review and meta-analysis

Margot H. Uijterwaal¹

Mariëlle Kocken¹

Anton. L.M. de Vries

Johannes Berkhof

Johannes C.F. Ket

Theo J.M. Helmerhorst

Chris J.L.M. Meijer

¹ Both authors contributed equally

Abstract

Objective Currently, women treated for high-grade cervical intraepithelial neoplasia (CIN 2/3) are followed-up by cytology to monitor them for residual and recurrent (post-treatment) disease. This systematic review and meta-analysis determine the test performance of testing for high-risk types of the human papillomavirus (hrHPV), cytology and co-testing (combined hrHPV testing and cytology) in predicting high-grade post-treatment disease (CIN2+).

Methods Studies that compared at least two of three post-treatment surveillance methods, and were published between January 2003 and May 2011, were identified through a bibliographic database search (PubMed, Embase.com and Wiley/Cochrane Library). Identification of relevant studies was conducted independently by two reviewers with a multi-step process. The reference standard used to diagnose post-treatment disease was histologically confirmed CIN2+. Sensitivity, specificity, diagnostic odds ratios and relative sensitivity and specificity were calculated for each study. Pooled estimates were calculated using a random effects model if heterogeneity among studies was significant, otherwise by using a fixed effects model. Estimates were reported with 95% confidence intervals (95%CI).

Results Out of 2410 potentially relevant citations, 8 publications, incorporating 1513 treated women, were included. Pooled sensitivities were 0.79 (95%CI 0.72-0.85) for cytology, 0.92 (0.87-0.96) for hrHPV testing, and 0.95 (0.91-0.98) for co-testing. HrHPV testing was more sensitive than cytology to predict post-treatment CIN2+ (relative sensitivity 1.15; 95%CI 1.06-1.25). Pooled specificities were 0.81 (95%CI 0.74-0.86) for cytology, 0.76 (0.67-0.84) for hrHPV testing and 0.67 (0.60-0.74) for co-testing. HrHPV testing and cytology had a similar specificity (relative specificity 0.95, 95%CI 0.88-1.02).

Discussion This review indicates that the hrHPV test should be included in post-treatment testing 6 months after treatment, because hrHPV testing has a higher sensitivity than cytology in detecting high-grade post-treatment disease and has a similar specificity.

Introduction

Women with high-grade cervical precursor lesions (Cervical Intraepithelial Neoplasia grade 2 (CIN2) and grade 3 (CIN3)) are treated by local excision or ablation to prevent progression to cervical cancer.¹⁻² Despite treatment, approximately 10.2% (95%CI 6.7-13.8)³ of these women are diagnosed with residual or recurrent (post-treatment) high-grade disease.³⁻⁹ Because of this substantial risk, many countries use surveillance strategies to identify post-treatment disease. These strategies fluctuate greatly between countries in both content, including follow-up modalities like cervical cytology, testing for high-risk types of the human papillomavirus (hrHPV) and colposcopy, either separately or in combination, and length of post-treatment surveillance.¹⁰⁻¹¹

Current Dutch national guidelines recommend cervical cytological testing at 6, 12 and 24 months after treatment, and, if necessary, annually for five years until three consecutive smears are read as normal.² After three consecutive negative smears, women return to five-yearly population-based routine screening. In contrast, treated women in the United Kingdom have a cytological examination at six and 12 months after treatment and, irrespective of the results, annual cytology for the subsequent nine years before reconsidering a return to routine screening.¹² In the United States annual cytology is even recommended for at least 20 years.¹³ In summary, treated women are followed more closely between two and 20 years before returning to population-based screening. Besides cytology, several other risk factors, including cone margin status, positive endocervical curettage and age, have been studied to predict recurrent cervical disease. However, these predictors are suboptimal.¹¹ For excision margins, for instance, previous studies, summarized in a review of Zielinski and colleagues, found that the sensitivity varied between 39 and 100%.⁹ Besides, this characteristic has shown to be less sensitive than cytology or hrHPV testing in predicting post-treatment disease.^{9,14}

More and more evidence is gathered concerning the use of hrHPV-testing during the follow-up period, because the presence of hrHPV is not only a prerequisite for the development of primary high-grade CIN¹⁵, but also for the development of post-treatment CIN.^{3,6,9,11,16-20} It is assumed that effective treatment not only removes the pre-malignant lesion, but also eliminates the responsible hrHPV infection.^{17,21} In women who develop post-treatment CIN the hrHPV infection stays present and is therefore associated with disease recurrence.^{11,17,21} The sensitivity of hrHPV-testing to detect post-treatment CIN outweighs that of cytology (approximately 90% versus 75%)^{3,9,11,16,20}, at the cost of a lower^{9,16} or similar^{3,19-20} specificity. Because this is not a screening setting, but post-treatment surveillance of a potentially lethal disease, sensitivity is valued higher than specificity. Strategies which include hrHPV testing are suggested as an alternative for conventional post-treatment surveillance.^{4,9-11,14,22} The majority of all treated women clear their hrHPV infection within 6 months²³ and have a significantly lower risk of

developing post-treatment CIN3+ than women without hrHPV clearance.^{4, 6, 11, 14, 22} In the surveillance of treated women combined testing with cytology and a hrHPV test (cotesting) results in a negative predictive value of over 99% to detect those at risk of developing post-treatment disease.^{6, 9, 14}

Previously published systematic reviews have examined the value of hrHPV testing in the context of post-treatment surveillance of CIN. All found a higher sensitivity for sole hrHPV testing or co-testing, compared to sole cytological testing.^{3, 9, 11, 16, 19-20} The review most recently published included studies up to 2007.¹⁶

We conducted this systematic review and meta-analysis to summarize and update current knowledge of the value of cytology, hrHPV testing and co-testing used in post-treatment surveillance. Besides describing the individual studies, we also determined the pooled sensitivity and specificity, diagnostic odds ratio and relative sensitivity and specificity.

Material and Methods

Search strategy

We searched the databases of PubMed, Embase.com, Wiley/Cochrane Library and WHO International Clinical Trials Registry for relevant studies published between January 2003 and April 2011. For this computer-aided search we used the following terms (including synonyms and closely related words) as index terms and free-text words: "vaginal smear" or "human papillomavirus" and "conisation" or "loop excision" or "CIN lesions" and "randomized controlled trials" or "systematic reviews". For the last two concepts we used predefined filters. These searches were not limited by language of publication. The example strategy for PubMed is presented in Supplementary Table 5.1 (S5.1). Previous systematic reviews or meta-analyses on the same subject as well as references of retrieved articles were used to search for additional relevant studies that could have been missed by the electronic search. Identification of relevant studies was conducted independently by two reviewers (MK and MU) with a multistep process (Figure 5.1). First, titles of the full list of citations were reviewed, followed by an assessment of abstracts of citations with potentially relevant titles. Disagreements were resolved by consensus. Finally, full-text articles of selected abstracts were considered for introduction in the review. We developed a data extraction sheet based on Cochrane guidelines²⁴⁻²⁵ to collect all relevant data from the studies, which was used by two reviewers (MK and MU). Disagreements were resolved after discussion. The following data were extracted: author, year and language of publication, country of study, population characteristics, study design, true positive (TP), false positive (FP), true negative (TN) and false negative (FN) values. The authors of papers which did not

state the values to construct a 2x2 table were contacted. Two independent reviewers (MK and MU) graded the methodological quality of the selected studies with a modified version of the QUality Assessment of Diagnostic Accuracy Studies (QUADAS) tool.²⁶ This modified version consists of 11 items on methodological characteristics that have the potential to introduce bias and are described in Supplementary Table 5.2 (S5.2). Items were scored positive (criteria satisfied), negative (criteria not satisfied), or unclear. We kept out two items (index test results blinded and relevant clinical information available) because hrHPV testing is performed by an objective test and is independent of clinical information. Furthermore we added the item "selection bias". Disagreements between the two extracting authors were resolved by consensus. Assessment of quality results was categorized.

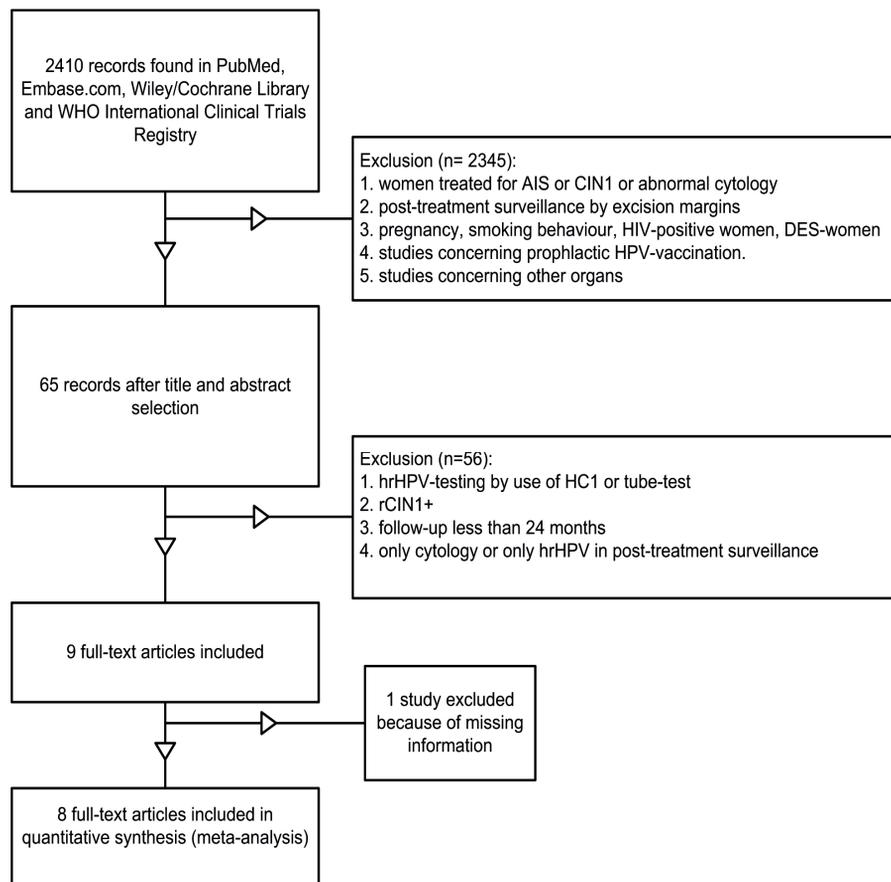


Figure 5.1 Study selection process for systematic reviews on accuracy of tests for detection of rCIN2+

Inclusion and exclusion criteria

To be included, both prospective and retrospective studies had to meet several inclusion criteria: (1) women should have been treated for CIN2/3 by either conisation (laser or coldknife) or LLETZ (Large Loop Excision of the Transformation Zone) procedure; (2) posttreatment surveillance should include at least two out of the following three methods; hrHPV testing, cytology, and/or co-testing (combined testing of cytology and hrHPV) at six months after treatment; (3) the positive endpoint, residual or recurrent high-grade post-treatment disease should be defined as a histological diagnosis of CIN2, CIN 3, adenocarcinoma in situ (AIS), adenocarcinoma, or squamous cell carcinoma (CIN2+). A negative endpoint should be defined as either a histological confirmation of no, or low-grade, disease (CIN0/1), or a repetitive negative cytological test result.

Both studies in which hrHPV testing was performed by Hybrid Capture II (HCII), as studies in which this was performed by the polymerase chain reaction (PCR) method were included. Both the sensitivities and negative predictive values of these tests to detect post-treatment CIN3+ seem similar, being 100% and 99% respectively²⁷ and show a good agreement.²⁸ In addition, we checked their similarity empirically by performing a bivariate regression analysis with type of hrHPV test as dichotomous covariate.

For a study to be considered for pooling, we required colposcopic evaluation of all positive test results in all women. Positive test results were defined as abnormal cytological test results, characterized as borderline dysplasia or worse, equivalent to atypical cells of unknown significance (ASC-US) or worse, and as hrHPV tests, positive for any hrHPV type.

Studies were excluded from this review, when they concerned the follow-up of women treated for adenocarcinoma in situ (AIS) or low-grade cervical disease (CIN1). Other exclusion criteria were studies concerning the follow-up of pregnant women, HIV-infected women, women exposed to diethylstilbestrol (DES) in utero, and studies concerning prophylactic HPV vaccination. Also studies with a follow-up of less than 12 months were discarded.

Statistical analysis

Outcome measures were the pooled estimates of sensitivity and specificity, the pooled estimate of the diagnostic odds ratio (DOR), and the pooled estimates of relative sensitivity and relative specificity. The DOR was defined as the odds of a positive test result in subjects with disease divided by the odds of a positive test result in subjects without disease.²⁹ The relative sensitivity was computed as the ratio of the sensitivity of the hrHPV test to the cytology test, and the relative specificity was computed analogously. In order to select the appropriate pooling method, the heterogeneity among the studies of each outcome measure was tested with Cochran's Q and quantified by I^2 .³⁰⁻³¹

If Cochran's Q was significant, the effects were pooled using a random effects model (REM)³², otherwise a fixed effects model (FEM) was used.

The sensitivity and specificity were pooled after applying the Freeman-Tukey double arcsine transformation³³⁻³⁴ and presented with a Clopper-Pearson exact 95% confidence interval (95%CI).³⁵ The 95%CI of the DOR was based on the standard error of the logarithm of the DOR. In all studies, cytology and hrHPV tests had been performed in all women. Therefore the standard error of the relative sensitivity and relative specificity was computed as the standard error of a relative risk of binary matched-pairs data.³⁶ A continuity correction of 0.5 was applied if the discordant cell frequencies equalled zero. The DOR as a measure of the discriminatory power of a test assumes that the thresholds between the different outcome categories remain constant over the included studies.³⁷ To check this assumption, the Pearson correlation between the logit true positive rate (TPR) and the logit false positive rate (FPR) was calculated for cytology, hrHPV testing and co-testing.³⁸ As a second check of heterogeneity across studies, Moses' regression model was fitted.³⁹ Finally, to check whether the type of hrHPV test used, or quality of the study differentially affected the sensitivity or specificity, a bivariate regression analysis of sensitivity and specificity was performed using maximum likelihood estimation with type of hrHPV test or study quality as covariates.⁴⁰

The sensitivity and specificity were pooled using R (version 2.13.0).⁴¹ The DOR was pooled, and the Pearson correlation between the logit TPR and logit FPR and Moses' regression model was fitted using Meta-DiSc (version 1.4).⁴² The relative sensitivity and relative specificity were pooled in Review Manager (version 5.1.2)⁴³, by importing the standard errors from a spreadsheet analysis in Microsoft Excel (2003). The bivariate regression analysis was performed in SAS (version 9.2). An effect with p-value <0.05 was considered statistically significant.

Results

The combined literature search identified 2410 citations. Of these, 2345 were excluded on title and abstracts, leaving 65 citations for full text review. Finally 9 citations met all criteria and were included in this review (Figure 5.1).^{4, 14, 28, 44-49} However, one study was later excluded, as the author did not respond to repetitive questioning about study design and results.²⁸

The eight studies remaining were heterogeneous in study characteristics, like study design, choice of hrHPV testing methods and the assessment of disease status at entry and end of follow-up (Table 5.1). Concerning design, seven studies were prospective, and one study was a case-control analysis. HrHPV testing was performed using HC2 testing in four studies, and PCR testing in three studies. In one study both techniques

Table 5.1 Study characteristics of the individual studies that investigated the performance of hrHPV and cytology (6 months after treatment) in predicting residual and recurrent high-grade disease.

Study	Year	Country	Participant final/initial	Recurrence* CIN2+ (%)	Follow-up in months (range)	hrHPV test	Study design
Cecchini ⁴⁴	2004	Italy	84/84	10 (11.9)	23 (11-40)	PCR	Prospective cohort study including 84 women (mean age 34.3) treated for CIN2/3 by LLETZ between February 1999 and June 2001. Follow-up at six months after treatment included hrHPV-testing, cytology and colposcopy. Method of PCR- testing: type-specific HPV6/E7 PCR. Method of cytology: not specified
Sarian ⁴⁷	2004	Brazil	88/107	11 (10.2)	17	HC2	Prospective cohort study including 88 women (mean age 34 years, range 20-60) treated for CIN2/3 by LLETZ between March 2001 and December 2002. Follow-up at six months after treatment for hrHPV-testing, cytology and colposcopy. Biopsies were taken if cytology revealed HSIL, or if a suspect area was present. Method of cytology: glass slide, specimen taken with Ayre spatula and endocervix brush.
Alonso ⁴⁴	2006	Spain	203/224	24 (11.8)	20 (6-66)	HC2	Prospective cohort study including 203 women (mean age 38.6 years, range 22-83) treated for CIN2/3 by LLETZ between May 1998 and October 2004. Follow-up at 6 months after treatment included hrHPV-testing, cytology and colposcopy. Biopsies were taken if abnormal cytology (\geq ASC-US) or hrHPV-positivity was present. Women with two consecutive negative cytological smears and negative colposcopy were considered negative for recurrence, irrespective of the hrHPV test result. Method of cytology: glass slide, specimen taken with Ayre spatula and cytobrush.
Kreimer ⁴	2006	USA	485/610	32 (6.6)	24	HC2 PCR	Prospective cohort study including 485 women (median age 24 years, range 21-28) treated for CIN2/3 by LLETZ between January 1997 and December 1998. Follow-up at 6 months after treatment included hrHPV-testing, including genotyping, and cytology. For analyses, data from HC2 testing was used and women with missing hrHPV test results were excluded. Method of PCR- testing: PGMY09/11 PCR. Method of cytology: liquid-based cytology.
Verguts ⁴⁹	2006	Belgium	72/72	6 (8.0)	24	HC2	Prospective cohort study including 72 women (mean age 40 years, range 22-78) treated for CIN2/3 by LLETZ between February 2000 and February 2003. Follow-up at three to six months after treatment included hrHPV, cytology and colposcopy. Biopsies were taken if any suspected area was present. Method of cytology: liquid-based cytology, taken with Cervexbrush.
Smart ⁴⁸	2010	New-Zealand	100/100	4 (4.0)	18	HC2	Prospective cohort study including 100 women (mean age 32 years, range 19-66), treated for CIN 2/3 by LLETZ or conization between January 2007 and January 2008. Follow-up (mean 9 months; range 3-18) included hrHPV-testing, cytology and colposcopy. Biopsies were taken if any suspect area was seen. One woman with inadequate cytology was excluded from analyses. Method of cytology: liquid-based cytology, taken with Cervexbrush.
Heymans ⁴⁵	2011	Belgium	63/63	n.a.	>24	PCR	Case control study (1:2) including 63 women (median age cases 40.9 years and controls 35.5 years) treated for CIN 2/3 by LLETZ or conization between January 2001 and December 2007. Follow-up at six months after treatment included hrHPV-testing, including genotyping, and cytology. Method of PCR- testing: type-specific HPV6/E7 PCR. Method of cytology: liquid-based cytology, taken with Cervexbrush.
Kocken ⁴⁶	2011	Netherlands	435/445	45 [†] (10.8)	24 [†]	PCR	Prospective cohort study including 435 women (mean age 34.9 years, range 21-70) treated for CIN2/3 by LLETZ or conization between July 1988 and November 2004. Follow-up at six months after treatment included hrHPV-testing, including genotyping, and cytology. Colposcopy was performed if abnormal cytology (\geq ASC-US) or hrHPV-positivity was present. For analyses, data was limited to 2 years of follow-up. Method of PCR- testing: GP5+/6+ PCR Method of cytology: partly glass slides, others not specified

CIN= Cervical intraepithelial neoplasia; hrHPV= high-risk types of the human papillomavirus; LLETZ= large loop excision of the transformation zone; PCR= polymerase chain reaction; HC2= Hybrid Capture 2 test; n.a.= not applicable

* Includes all residual and recurrent disease, [†] For analyses data was limited to 2 years of follow-up.

were performed. All studies collected samples for cytology and hrHPV-testing six months after treatment. One study had a significant longer follow-up than all other included studies.⁴⁶ Therefore, data for this study was recalculated for a follow-up of 2 years. In all studies combined, 1513 women had been treated for CIN2/3. The number of women per study varied between 63 and 485. The age of the participants ranged between 19 and 83 years. Treatment failure, expressed in terms of residual and recurrent CIN2+, ranged from 4.0% to 11.9%. Recurrence rates of CIN2+, sensitivities as well as DORs of the individual studies and pooled values at six months post treatment are shown in Figure 5.2. The correlation between the logit TPR and logit FPR was not significant for all diagnostic tests (cytology $r=0.024$, $p=0.96$; hrHPV $r = 0.49$, $p = 0.18$; and co-testing: $r = 0.42$, $p = 0.26$). Moses' linear regression model did not significantly improve the constant model (cytology: $p = 0.21$, hrHPV: $p = 0.80$, and co-testing: $p = 0.67$), giving further support for the assumption of a constant DOR over studies.

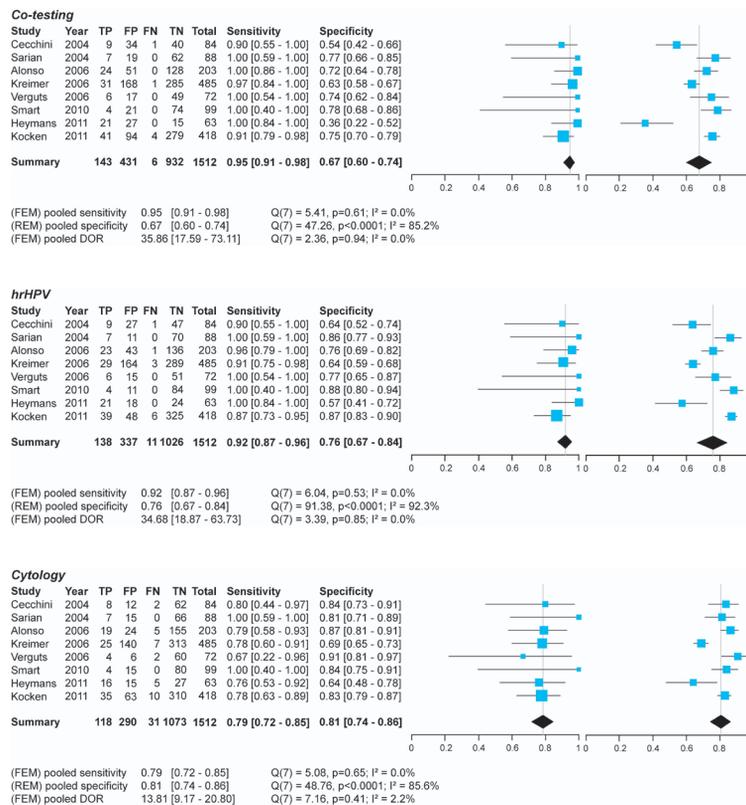


Figure 5.2 Meta-analysis of the sensitivity and specificity, including pooled estimates of testing 6 months after treatment with cytology, hrHPV or co-testing. Forest plots of sensitivity (left) and specificity (right).

Abbreviations: TP=true positives, FN= false negatives, TN= true negatives, FEM= fixed effects model, REM= random effects model, DOR= diagnostic odds ratio.



Diagnostic accuracy

The sensitivity of cytological testing 6 months after treatment in predicting post treatment CIN lesions varied between the studies between 0.67 and 1.0. For sensitivity, the test for heterogeneity between studies was not significant ($Q(7)=5.08$, $p=0.65$; $I^2=0.0\%$). The pooled sensitivity was 0.79 (95%CI 0.72-0.85). The specificity of cytology ranged from 0.64 to 0.91 and heterogeneity between studies was significant ($Q(7)=48.76$, $p<0.0001$; $I^2=85.6\%$). The pooled specificity was 0.81 (95%CI 0.74-0.86). The pooled DOR of cytology was 13.81 (95%CI 9.17-20.80) and there was no evidence for statistically significant heterogeneity ($Q(7)=7.16$, $p=0.41$; $I^2=2.2\%$).

In the studies the sensitivity of hrHPV-testing varied between 0.87 and 1.0. For sensitivity, the test for heterogeneity between studies was not significant ($Q(7)=6.04$, $p=0.53$; $I^2=0.0\%$). The pooled sensitivity was 0.92 (95%CI 0.87-0.96). The specificity of hrHPV ranged from 0.57 to 0.88 and heterogeneity between studies was significant ($Q(7)=91.38$, $p<0.0001$; $I^2=92.3\%$). The pooled specificity was 0.76 (95%CI 0.67-0.84). The pooled DOR of hrHPV testing was 34.68 (95%CI 18.87-63.73) and there was no evidence for statistically significant heterogeneity ($Q(7)=3.39$, $p=0.85$; $I^2=0.0\%$).

The sensitivity of co-testing varied between 0.90 and 1.0. For sensitivity, the test for heterogeneity between studies was not significant ($Q(7)=5.41$, $p=0.61$; $I^2=0.0\%$). The pooled sensitivity was 0.95 (95%CI 0.91-0.98). The specificity varied between 0.36 and 0.78 and heterogeneity between studies was significant ($Q(7)=47.26$, $p<0.0001$; $I^2=85.2\%$). The pooled specificity was 0.67 (95%CI 0.60-0.74). The pooled DOR of co-testing was 35.86 (95%CI 17.59-73.11) and there was no evidence for statistically significant heterogeneity ($Q(7)=2.36$, $p=0.94$; $I^2=0.0\%$).

A bivariate regression model was fitted to test the influence of the two types of hrHPV tests, HC2 and PCR on sensitivity and specificity. Adding of the hrHPV test type as a covariate explained 11.5% of the between-study variance for cytology, which was not significant ($t(7) = 1.25$, $p= 0.25$). Also for hrHPV and co-testing the addition of the hrHPV test type as a covariate explained a not significant part of the between-study variance (hrHPV 7.7%, ($t(7) = 0.81$, $p= 0.45$) and co-testing 31.9%, ($t(7) = 2.15$, $p= 0.07$)).

Overall, hrHPV-testing after six months predicted post-treatment CIN with significantly higher sensitivity (relative sensitivity 1.15; 95%CI 1.06-1.25 ($Z=3.27$, $p=0.001$)) than cytology and a similar specificity (relative specificity 0.95; 0.88-1.02 ($Z=1.53$, $p=0.13$)).

Methodological quality of included studies

Table 5.2 summarizes the methodological quality of the 8 included studies. All studies included a representative patient spectrum of women treated for high-grade cervical disease. One case-control study was included and selection bias could not be excluded.⁴⁵ We added this criterion and although most studies imply that women were continuously

included, this was only explicitly mentioned in one study.⁴⁴ The reference standard was adequate in all studies. All studies defined a positive test result of post-treatment disease as a histological finding of CIN2+. A negative test result was verified by colposcopy in five studies. In the three studies remaining a negative test result was verified by consecutive negative cytological smears.^{14, 45-46} One of these studies⁴⁶ was a multi-cohort study and in two of the three incorporated studies test results were confirmed with a colposcopic examination in all women, and in the third study absence of disease was confirmed by three consecutive negative cytological smears¹⁷. In every study all patients were assessed within two years. Complete verification with the reference standard was performed in four studies.^{4, 14, 47, 49} In three other studies not all patients were examined by colposcopy, but were considered free of disease by cytological examination.^{14, 45-46} One study performed a colposcopy in all patients at 6 months after treatment, but did not specify when colposcopies were performed later in follow-up.⁴⁸ In none of the studies hrHPV was part of the reference standard. Moreover, in most studies the reference standards were interpreted without knowing the results of the hrHPV-test. Only in two studies women were referred on basis of the hrHPV test result.^{14, 46} For the case-control study this item could not be assessed.⁴⁵ Of interest might be that only one study explicitly mentions biopsy taking in all women irrespective of any test result⁴, in all other studies biopsy taking is dependent of either visual impression^{14, 47-49} or abnormal cytological test results.^{14, 44, 46-47} If present, uninterpretable results were mentioned in all studies. Although in one study the percentage "missing" was higher in the group without post-treatment disease than in the group diagnosed with posttreatment CIN2+ (13.7% versus 5.9%), this difference was not significant ($p=0.295$, Fisher's exact).⁴ Half of the studies had no withdrawals to explain and the remaining studies dealt with this item appropriately.

The included studies differed in only 3 items of the QUADAS list (Table 5.2); selection bias, differential verification and blinding of reference test results. These three items were each separately added as a dichotomous covariate to the bivariate regression models that were also used to test the effect of the hrHPV test type. A rating of '+' was counted as present, a rating of '?' or '-' as absent. Separate addition of each item as a covariate did not improve the bivariate regression models of cytology, hrHPV-testing or co-testing significantly (data not shown).

Discussion

In this systematic review we described the value of 6-month testing for cytology and/or hrHPV in the surveillance of women treated for CIN2/3 and confirmed the advantage of implementing hrHPV in the follow-up of women treated for high-grade CIN as found

Table 5.2 Summary of methodological quality. Review authors' judgments about each methodological quality item for each included study.

	Kocken 2011 ⁴⁶	Heymans 2011 ⁴⁵	Smart 2010 ⁴⁸	Alonso 2006 ¹⁴	Kreimer 2006 ⁴	Verguts 2006 ⁴⁹	Sarian 2004 ⁴⁷	Cecchini 2004 ⁴⁴
Representative spectrum?	+	+	+	+	+	+	+	+
Selection bias	?	-	?	?	?	?	?	+
Acceptable reference standard?	+	+	+	+	+	+	+	+
Appropriate timing of tests and verification of outcome?	+	+	+	+	+	+	+	+
Partial verification avoided?	+	+	+	+	+	+	+	+
Differential verification avoided?	-	-	?	-	+	+	+	+
Incorporation avoided?	+	+	+	+	+	+	+	+
Reference test results blinded?	-	?	+	+	+	+	+	+
Uninterpretable results reported?	+	+	+	+	+	+	+	+
Withdrawals explained?	+	+	+	+	+	+	+	+

Minus sign= negative score, plus sign= positive score, question mark= unclear whether item scores negatively or positively

in previously conducted meta-analyses.^{3, 9, 11, 16, 19-20} HrHPV testing has a significantly higher sensitivity than cytology, indicated by a relative sensitivity of 1.15 (95%CI 1.06-1.25), without decreasing the specificity (relative specificity 0.95, 0.88-1.02).

We measured the DOR to compare the three different tests (hrHPV, cytology and co-testing). The DOR of co-testing testing was higher than the DOR of hrHPV-testing or cytology. This indicates that the overall discriminative power of co-testing is the best. As approximately 10%³ of women treated for CIN2/3 develop high grade post-treatment disease, it seems logical to choose a test that assures a minimal risk of high-grade disease in this group and to select the test with the highest sensitivity. The pooled sensitivity of co-testing was with 0.95 (95%CI 0.91-0.98) also higher than the sensitivity of sole cytological or hrHPV testing. However, the pooled specificity was only 0.67, resulting in approximately 10% more women referred for repeat testing or colposcopy, or both.

Sources of bias and potential sources of heterogeneity

The purposes of a quality assessment are to identify potential sources of bias and to estimate their impact. The overall methodological quality of the included studies was generally good (Table 5.2), however study characteristics between studies varied. For instance, although the average age in most studies was approximately 35 years, one study included women with a median age of 24 years (range 21-28).⁴ As both the prevalence of CIN and of hrHPV varies with age, this factor may influence the test accuracy

across the studies. Another difference was that percentages of recurrence were given for the total follow-up period. Most studies measured follow-up until 24 months after treatment and by limiting the follow-up of Kocken et al. to two years⁴⁶, studies became more homogeneous. Another point to address is that we considered the first hrHPV testing moment to be at 6 months after treatment. Yet, some studies performed hrHPV tests before^{4, 47, 49} or after^{14, 48} 6 months. Because hrHPV infections clear over time^{17, 21}, the risk of developing post-treatment disease will also diminish over time and this could possibly affect the sensitivity of the test.

Other possible reasons for heterogeneity between the studies may be explained by different collection methods of material (e.g. liquid based cytology versus conventional glass slides) and different execution of the analyses (e.g. either single pathologist or review of all cytology and/or histology).

Based on our methodological appraisal the most likely sources of bias are selection bias and differential verification bias. Selection bias arises when women are not included in a consecutive order. Only one of the cohort studies⁴⁴ mentions the inclusion of patients to be explicitly consecutive, the other studies only describe to include women within a certain timeframe. However, the populations included in these studies seem to be consecutive and are most likely comparable to one another. But, as for instance, this would have resulted in an exclusion of more difficult cases, it could have resulted in a lower number of false positives and false negatives and hence to increased estimates of sensitivity and specificity.

Differential verification bias arises if two different reference tests are used and the tests have different accuracy. Some of the studies included in our analysis avoided this problem because they performed colposcopic examinations in all women, irrespective of their test results.^{4, 44, 47, 49} However, only one study has, besides performing a colposcopy, also taken a biopsy in all subjects to verify the presence, or absence, of cervical disease. Other studies only referred for exit-colposcopy when abnormal test results were found.^{14, 46} Some studies did not specify when women were referred for colposcopy.⁴⁸ These last studies are therefore prone to (detection) bias. Differential verification bias could have resulted in an overestimation of both sensitivity and specificity. However, women with repetitively negative cytological test results have a low risk of harbouring high-grade disease⁴⁶ and therefore we expect this type of bias to be of limited influence. The included studies remained statistically heterogeneous concerning specificity. Consequently we used a random effect model to calculate the pooled estimates that were heterogeneous. Until better-conducted studies, such as large randomized controlled trials with histological verification of all subjects, independent of hrHPV or cytology results, are available, the pooled estimate provides clinically relevant information.

Study limitations

Our study provides evidence suggesting that hrHPV-testing is more accurate for the diagnosis of post-treatment cervical disease than cytology. Although this is in line with previous reviews, these results are based on a small number of studies.

Also our review is limited to a follow-up of two years as almost all included studies had a follow-up restricted to this period and thereby hampering long-term information of the different test performances. This information is relevant, as the risk of developing recurrent disease is significantly increased for over 10 years after treatment.^{5, 7, 18} Only two studies were identified describing hrHPV testing in long-term follow-up.^{46, 50} One of these studies excluded residual/recurrent disease developed within two years of treatment and only described the performance of the hrHPV test. Therefore no comparison could be made with the performance of cytology and this study was excluded for analysis in this review.⁵⁰ One study that described the long-term prediction of hrHPV and/or cytology at 6 months after treatment remained.⁴⁶

In this study the sensitivities of hrHPV and cytology in detecting post-treatment disease decreased when the total follow-up time increased. For cytology, the sensitivity in the total study with a follow-up up to 21.5 years was 66% (95%CI 55-76) compared to a sensitivity of 78% in the first two years (Figure 5.2). For hrHPV testing, these values were 72% (95%CI 61-81) and 87%, respectively. This could illustrate the acquisition of new hrHPV types or re-infection/ reactivation of the same hrHPV types that will eventually result in high-grade lesions. Another possibility might be the presence of false negative HPV tests, due to integration of the viral DNA targeted in the HPV test in the genome of the host cell.

Another limitation is that only 6-month testing is measured. Several studies indicated that one test moment is insufficient to predict post-treatment disease and that therefore repeat testing should be performed.^{46, 50} However, although some authors have described follow-up algorithms for post-treatment surveillance, pooling of these data was not possible.

A final limitation is that although a large number of citations were reviewed, this review might be subject to publication bias. However, the impact that publication bias has on diagnostic accuracy systematic reviews is unknown.⁵¹

Conclusions

This review clearly indicates that post-treatment testing at 6 months after treatment should include hrHPV testing. HrHPV testing after 6 months has a higher sensitivity than and a similar specificity as cytology. The sensitivity of co-testing is even higher than of the separate individual tests. As women treated for CIN2/3 have a high risk of

developing recurrent disease, sensitivity is valued higher than specificity and therefore hrHPV testing (or co-testing) should be incorporated in post-treatment surveillance. As even the sensitivity of co-testing is not sufficiently high to rely on a single test moment, repeat testing is necessary to identify all women at risk for post-treatment disease.⁴⁶ Several studies already indicated that women testing negative for co-testing after 6 months could omit the 12-month test moment and return for monitoring at 24 months after treatment. However, more information, especially on long-term recurrence and cost-effectiveness is needed to recommend a definite follow-up algorithm.

References

- Martin-Hirsch PP, Paraskevaïdis E, Bryant A, Dickinson HO, Keep SL. Surgery for cervical intraepithelial neoplasia. *Cochrane Database Syst Rev*. 2010;6:CD001318.
- NVOG. National Guideline "Cervical Intraepithelial Neoplasia". [Webpage] 2004. Available from: http://www.oncoline.nl/richtlijn/item/pagina.php?richtlijn_id=220. Cited September 28th, 2011.
- 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.
- Kreimer AR, Guido RS, Solomon D, Schiffman M, Wacholder S, Jeronimo J, 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(5):908-14.
- 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. *JNCI*. 2009;101(10):721-8.
- Prato B, Ghelardi A, Gadducci A, Marchetti I, Di Cristofano C, Di Coscio G, 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(1):90-4.
- Soutter WP, Sasieni P, Panoskaltis T. Long-term risk of invasive cervical cancer after treatment of squamous cervical intraepithelial neoplasia. *Int J Cancer*. 2006;118(8):2048-55.
- Wright TC. Precancerous lesions of the cervix. In: Kumar RJ, editor. *Blaustein's pathology of the female genital tract*. 4th ed. New York: Springer Verlag; 1995. p.248-57.
- Zielinski GD, Bais AG, Helmerhorst TJ, Verheijen RH, de Schipper FA, Snijders PJ, et al. HPV testing and monitoring of women after treatment of CIN 3: review of the literature and meta-analysis. *Obstet Gynecol Surv*. 2004;59(7):543-53.
- Kyrgiou M, Tsoompou I, Vrekoussis T, Martin-Hirsch P, Arbyn M, Prendiville W, et al. The up-to-date evidence on colposcopy practice and treatment of cervical intraepithelial neoplasia: the Cochrane colposcopy & cervical cytopathology collaborative group (C5 group) approach. *Cancer Treat Rev*. 2006;32(7):516-23.
- Paraskevaïdis E, Arbyn M, Sotiiriadis A, Diakomanolis E, Martin-Hirsch P, Koliopoulos G, 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(2):205-11.
- Colposcopy and Programme Management. Guidelines for the NHS Cervical Screening Programme. In: Luesley DLS, editor. *NHS cancer screening programmes*. Second ed. Sheffield: NHS Cancer Screening Programmes; 2010.
- Bulletins-Gynecology ACoP. ACOG Practice Bulletin no. 109: Cervical cytology screening. *Obstet Gynecol*. 2009;114(6):1409-20.
- Alonso I, Torne A, Puig-Tintore LM, Esteve R, Quinto L, Campo E, et al. Pre- and post-conization highrisk HPV testing predicts residual/recurrent disease in patients treated for CIN 2-3. *Gynecol Oncol*. 2006;103(2):631-6.
- 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(1):12-9.
- Chan BK, Melnikow J, Slee CA, Arellanes R, Sawaya GF. Posttreatment human papillomavirus testing for recurrent cervical intraepithelial neoplasia: a systematic review. *Am J Obstet Gynecol*. 2009;200(4):422 e1-9.
- Nobbenhuis MA, Meijer CJ, van den Brule AJ, Rozendaal L, Voorhorst FJ, Risse EK, 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(6):796-801.
- Strander B, Andersson-Ellstrom A, Milsom I, Sparren P. Long term risk of invasive cancer after treatment for cervical intraepithelial neoplasia grade 3: population based cohort study. *BMJ*. 2007;335(7629):1077.
- Arbyn M, Paraskevaïdis 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(3 Suppl 1):S7-11.
- 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.
- Elfgren K, Jacobs M, Walboomers JM, Meijer CJ, Dillner J. Rate of human papillomavirus clearance after treatment of cervical intraepithelial neoplasia. *Obstet Gynecol*. 2002;100:965-71.
- Kitchener HC, Walker PG, Nelson L, Hadwin R, Patnick J, Anthony GB, et al. HPV testing as an adjunct to cytology in the follow up of women treated for cervical intraepithelial neoplasia. *BJOG*. 2008;115(8):1001-7.
- Nobbenhuis MA, Helmerhorst TJ, van den Brule AJ, Rozendaal L, Voorhorst FJ, Bezemer PD, et al. Cytological regression and clearance of high-risk human papillomavirus in women with an abnormal cervical smear. *Lancet*. 2001;358(9295):1782-3.
- Higgins JPT GS. *Cochrane Handbook for Systematic Reviews of Interventions*. In: *The Cochrane Collaboration*, editor. Version 5.1.0 ed; 2011.
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.
- Reitsma JB RA, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Assessing methodological quality. In: Deeks JJ, PM B, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy: The Cochrane Collaboration*; 2009.

27. Soderlund-Strand A, Rymark P, Andersson P, Dillner J, Dillner L. Comparison between the Hybrid Capture II test and a PCR-based human papillomavirus detection method for diagnosis and post-treatment follow-up of cervical intraepithelial neoplasia. *J Clin Microbiol.* 2005;43(7):3260-6.
28. Kang WD, Oh MJ, Kim SM, Nam JH, Park CS, Choi HS. 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):e1-6.
29. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol.* 2003;56(11):1129-35.
30. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21(11):1539-58.
31. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-60.
32. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177-88.
33. Freeman MF, Tukey JW. Transformations related to the angular and the square root. *Ann Math Stat.* 1950;21(4):607-11.
34. Miller JJ. The inverse of the Freeman-Tukey double arcsine transformation. *Am Stat.* 1978;32(4):138.
35. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika.* 1934;26(4):404-13.
36. Agresti A, Min Y. Effects and non-effects of paired identical observations in comparing proportions with binary matched-pairs data. *Stat Med.* 2004;23(1):65-75.
37. Irwig L, Bossuyt P, Glasziou P, Gatsonis C, Lijmer J. Designing studies to ensure that estimates of test accuracy are transferable. *BMJ.* 2002;324(7338):669-71.
38. Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol.* 2002;2:9.
39. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med.* 1993;12(14):1293-316.
40. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med.* 2002;21(4):589-624.
41. R Development Core team. R: A Language and Environment for statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2010.
42. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol.* 2006;6:31.
43. The Cochrane Collaboration. Review Manager (RevMan). 5.1.2 ed. Copenhagen: the Nordic Cochrane Centre, The Cochrane collaboration; 2008.
44. 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(2):225-8.
45. Heymans J, Benoy IH, Poppe W, Depuydt CE. Type-specific HPV genotyping improves detection of recurrent high-grade cervical neoplasia after conisation. *Int J Cancer.* 2011;129(4):903-9.
46. 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(5):441-50.
47. Sarian LO, Derchain SF, Andrade LA, Tambascia J, Morais SS, Syrjanen KJ. HPV DNA test and Pap smear in detection of residual and recurrent disease following loop electrosurgical excision procedure of highgrade cervical intraepithelial neoplasia. *Gynecol Oncol.* 2004;94(1):181-6.
48. 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(2):164-7.
49. Verguts J, Bronselaer B, Donders G, Arbyn M, Van Eldere J, Drijkoningen M, 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(11):1303-7.
50. Strander B, Ryd W, Wallin KL, Warleby B, Zheng B, Milsom I, 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(12):1849-55.
51. De Vet HC, Eisinga A, Riphagen II, Aertgeers B, Pewsner D. Searching for studies. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy: The Cochrane collaboration*; 2008.

Supplementary Table 5.1 (S5.1) Search strategy using PubMed (April 6th, 2011)

(“Vaginal Smears”[Mesh] OR “Papillomaviridae”[Mesh] OR “DNA Methylation”[Mesh] OR “Biopsy”[Mesh:NoExp] OR “Colposcopy”[Mesh] OR methylation*[tiab] OR colposcop*[tiab] OR smear*[tiab] OR papanicolaou[tiab] OR hpv[tiab] OR papillomavir*[tiab] OR biops*[tiab]) AND (conizati*[tiab] OR conisati*[tiab] OR lletz[tiab] OR letz[tiab] OR leep[tiab] OR loop[tiab] AND (excision*[tiab] OR electroexcision*[tiab])) OR “Conization”[Mesh] OR “Cervical Intraepithelial Neoplasia”[Mesh] OR cin[tiab] OR “cervical intraepithelial”[tiab] OR “Uterine Cervical Dysplasia”[Mesh] OR ((cervical[tiab] OR cervix[tiab]) AND dysplas*[tiab])) AND ((randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab]) OR ((review*[tiab] OR search*[tiab] OR survey*[tiab] OR handsearch*[tiab] OR handsearch*[tiab]) AND (databa*[tiab] OR data-ba*[tiab] OR bibliograph*[tiab] OR electronic*[tiab] OR medline*[tiab] OR pubmed*[tiab] OR embase*[tiab] OR Cochrane[tiab] OR cinahl[tiab] OR psycinfo[tiab] OR psychinfo[tiab] OR cinhal[tiab] OR “web of science”[tiab] OR “web of knowledge”[tiab] OR ebSCO[tiab] OR ovid[tiab] OR mrct[tiab] OR metaregist*[tiab] OR meta-regist*[tiab] OR ((predetermined[tiab] OR pre-determined[tiab]) AND criteri*[tiab]) OR apprais*[tiab] OR inclusion criteri*[tiab] OR exclusion criteri*[tiab]) OR (review[pt] AND systemat*[tiab]) OR “systematic review”[tiab] OR “systematic literature”[tiab] OR “integrative review”[tiab] OR “integrative literature”[tiab] OR “evidence-based review”[tiab] OR “evidence-based overview”[tiab] OR “evidence-based literature”[tiab] OR “evidence-based survey”[tiab] OR “literature search”[tiab] OR ((systemat*[ti] OR evidencebased[ti]) AND (review*[ti] OR literature[ti] OR overview[ti] OR survey[ti])) OR “data synthesis”[tiab] OR “evidence synthesis”[tiab] OR “data extraction”[tiab] OR “data source”[tiab] OR “data sources”[tiab] OR “study selection”[tiab] OR “methodological quality”[tiab] OR “methodologic quality”[tiab] OR cochrane database syst rev[ta] OR meta-analy*[tiab] OR metaanaly*[tiab] OR metanaly*[tiab] OR meta-analysis[pt] OR metasynthesis[tiab] OR metasynthesis[tiab] OR meta-study[tiab] OR metastudy[tiab] OR metaethnograph*[tiab] OR meta-ethnograph*[tiab] OR Technology Assessment, Biomedical[mh] OR hta[tiab] OR health technol assess [ta] OR evid rep technol assess summ[ta] OR health technology assessment[tiab]))

Supplementary Table 5.2 (S5.2): Checklist for the Quality Assessment of Diagnostic Accuracy Studies (Quadas), modified version.²⁶

nr	Item definition	Item question	Assessment
1	Representative spectrum?	Was the spectrum of included patients representative of the patients who will receive the test in practice?	Yes: patients treated by conisation (either LLETZ or cone) for cervical intraepithelial neoplasia grade 2 or 3 (CIN2/3) Unclear: reporting insufficient to assess this item No: inclusion of healthy controls
2	Selection bias avoided?	Were included women included on a continuous basis?	Yes: women were explicitly included on a continuous basis Unclear: continuous selection is not explicitly mentioned No: women were not continuously included (e.g. case control study)
3	Acceptable reference standard?	Is the reference standard likely to correctly classify the target condition?	Yes: histological confirmation of CIN2+ is considered as presence of disease, absence of disease is either confirmed by colposcopy or by repetitive negative cytology results Unclear: reporting insufficient to assess this item No: no histological confirmation of presence of disease or no repetitive negative cytology results
4	Acceptable timing of tests and verification of outcome?	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	'Yes' for all studies because all studies assessed disease status within 2 years of follow-up.
5	Partial verification avoided?	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	'Yes' for all studies as all patients performed either a colposcopy in all patients or were considered negative for disease because of ≥ 2 cytological negative smears:
6	Differential verification avoided?	Did all patients receive the same reference standard regardless or do some the index test result?	Yes: irrespective of the index test result the same reference standard is performed Unclear: reporting insufficient to assess this item No: the result of the index test effected the policy of verification (e.g. a positive hrHPV test results in colposcopy) or different reference standards were performed.
7	Incorporation avoided?	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes: the hrHPV test is not part of the reference standard Unclear: reporting insufficient to assess this item No: the hrHPV test is part of the reference standard
8	Index test results blinded?	Were the index test results interpreted without knowledge of the results of the reference standard?	Not applicable. The performed hrHPV tests (both HCII and PCR) are objective tests.
9	Reference test results blinded?	Were the reference test results interpreted without knowledge of the results of the index test?	Yes: hrHPV test was performed prior to or simultaneously with reference test of hrHPV test results were blinded Unclear: reporting insufficient to assess this item No: hrHPV test was performed and assessed with knowledge of the results of the reference standard
10	Clinical data available?	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not applicable. The performance of the(automated) hrHPV test is not influenced by the availability of clinical data
11	Uninterpretable results reported?	Were uninterpretable or intermediate test results reported?	Yes: the number of patients with indeterminate hrHPV test results has been reported Unclear: reporting insufficient to assess this item No: the number of patients with indeterminate hrHPV test results has not been reported
12	Withdrawals explained?	Were withdrawals from the study explained?	Yes: the number of withdrawals has been reported and reasons have been explained or there were no withdrawals Unclear: reporting insufficient to assess this item No: the number of withdrawals has not been reported