

BMJ Open Comparative effectiveness study on human papillomavirus detection methods used in the cervical cancer screening programme

Mari Nygård,¹ Kjetil Røysland,¹ Suzanne Campbell,¹ Joakim Dillner²

To cite: Nygård M, Røysland K, Campbell S, *et al.* Comparative effectiveness study on human papillomavirus detection methods used in the cervical cancer screening programme. *BMJ Open* 2013;**4**:e003460. doi:10.1136/bmjopen-2013-003460

► Prepublication history and additional material for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2013-003460>).

Received 21 June 2013
Revised 15 November 2013
Accepted 29 November 2013



CrossMark

¹Department of Research, Cancer Registry of Norway, Oslo, Norway

²Department of Laboratory Medicine and the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Correspondence to

Dr Mari Nygård; mari.nygard@kreftregisteret.no

ABSTRACT

Objectives: To compare the short-term and long-term effectiveness of human papillomavirus (HPV) tests in Norwegian Cervical Cancer Screening Programme (NCCSP).

Design: Nationwide register-based prospective follow-up study.

Setting: In 2005, the NCCSP implemented HPV testing in follow-up of unsatisfactory, atypical squamous cells of undetermined significance (ASC-US) and low-grade squamous intraepithelial lesion (LSIL) cytology.

Participants: 19 065 women with repeat cytology and HPV test after unsatisfactory ASC-US or LSIL screening result in 2005–2009.

Interventions: Through individual registry linkages we observed how women were treated in the regular medical care.

Main outcome measures: We estimated cumulative incidence of cervical intraepithelial neoplasia grade 2 or worse (CIN2+) in 6 months and 3 years after repeat cytology and HPV test. Patients diagnosed with CIN2+ in 6 months and 3 years were assessed for initial HPV positivity.

Results: 5392 had ASC-US/LSIL and 13 673 had normal/unsatisfactory repeat cytology; for HPV detection 4715 used AMPLICOR HPV Test (Roche Diagnostics, Basel, Switzerland), 9162 Hybrid Capture 2 (HC2) High-Risk HPV DNA Test (QIAGEN, Gaithersburg, Maryland, USA) and 5188 PreTect HPV-Proofer (NorChip, Klokkarstua, Norway). Among those with ASC-US/LSIL repeat cytology, 3-year risk of CIN2+ was 15-fold in Amplicor/HC2-positives compared with Amplicor/HC2-negatives and sevenfold in Proofer-positives compared with Proofer-negatives; a 3-year risk of CIN2+ was 2.1% (95% CI 0.7% to 3.4%) in Amplicor-negatives and 7.2% (95% CI 5.4% to 8.9%) in Proofer-negatives. Close to 100% of patients with CIN2+ diagnosed within 6 months tested positive to HPV (all methods). Considering all patients diagnosed with CIN2+ in 3-year follow-up, 97% were initially positive in the Amplicor group and more than 94% in the HC2 group, compared with less than 80% in the Proofer group.

Conclusions: While the long-term evaluation of new screening routines showed a good overall performance of triage-HPV DNA testing, the management of HPV-negative women with persistent ASC-US/LSIL was suboptimal.

Strengths and limitations of this study

- Using the comprehensive Norwegian registry for human papillomavirus (HPV) tests, smears and histopathologies, we were able to demonstrate the value of the three most commonly used HPV tests in diagnosing precancerous lesion in the cervix uteri, when a HPV test was used as part of routine care.
- Results from this study are directly applicable in countries with comparable screening recommendations.
- Long-term and comprehensive comparative effectiveness studies of newly proposed technology are needed in order to deliver the best possible healthcare.

INTRODUCTION

Cytology-based cervical cancer screening programmes reduce the incidence of and mortality from cervical cancer^{1–2} by identifying and treating women with asymptomatic preinvasive cervical intraepithelial neoplasia grade 2 or 3 (CIN2 or 3). Following the establishment of the causal association between high-risk human papillomavirus (hrHPV) infection and cervical cancer,³ HPV testing represents a new technology that can offer more effective screening schemes.^{4–5}

In Norway, a nationwide cytology-based organised cervical screening programme, the Norwegian Cervical Cancer Screening Programme (NCCSP), started in 1995. Approximately 430 000 cervical cytological smears are taken annually, of which only 1.4% show severe abnormalities, such as atypical squamous cells, cannot rule out a high-grade lesion (ASC-H), high-grade squamous intraepithelial lesion (HSIL), adenocarcinoma in situ (AIS) and cervical cancer, that require immediate follow-up with colposcopy and biopsy. About 7% of screening cytology results are unsatisfactory, atypical squamous cells of undetermined significance (ASC-US) or low-grade squamous intraepithelial lesion (LSIL),

and these require additional investigation for proper clinical management. In accordance with the European consensus for triage options,⁶ repeat cytology after 6 months is applied.⁷ If repeat cytology does not show severe abnormalities, further clinical follow-up is determined from cytology in combination with simultaneous HPV test result.^{6,8}

To evaluate the effectiveness of the NCCSP policy, including the different HPV detection methods actually used, women triaged with HPV testing in Norway were identified and followed up for the development of CIN2 or worse (CIN2+).

MATERIAL AND METHOD

Cervical cancer screening in Norway

The Cancer Registry of Norway receives reports of cancer and precancerous lesions since 1953 and administrates the NCCSP since 1995.^{9,10} Compulsory central registration of cervical cytology results, started in November 1991, is estimated to be close to 100% complete as all laboratories, including private ones, are required by law to report results for cervical cytology, histology (from 2002) and HPV test (from 2005).¹¹ Each Regional Health Trust authority decided which of the commercially available HPV methods should be used in their area, and as a result, different HPV methods have been used in different geographical areas in Norway.

Conventional cytology was diagnosed and registered according to the Bethesda system as negative for intraepithelial lesions or malignancy (NILM), unsatisfactory, ASC-US, LSIL, ASC-H, HSIL, AIS and cervical cancer.¹² Histology diagnoses are registered according to the guidelines of the WHO as CIN2+, CIN3 or worse (CIN3+) or cervical cancer.¹³

NCCSP recommends women 25–69 years of age to take a cytology smear every 3 years. For those with unsatisfactory, ASC-US and LSIL, the policy included a repeat cytology with HPV testing in 6 months (figure 1A). If repeat cytology showed severe abnormalities, or if it showed ASC-US or LSIL in combination with a positive triage-HPV test result, colposcopy and biopsy were recommended. Those with a negative HPV result combined with ASC-US/LSIL, unsatisfactory or NILM repeat cytology was returned to the regular 3-year screening programme. For women with NILM or unsatisfactory repeat cytology combined with a positive HPV test result, an additional follow-up HPV test was recommended in 6–12 months.

The three most common HPV detection methods during the study period (2005–2010) were AMPLICOR HPV Test (Roche Diagnostics, Basel, Switzerland), Hybrid Capture 2 (HC2) High-Risk HPV DNA Test (QIAGEN, Gaithersburg, Maryland, USA) and Pre-tect HPV-Proofer (NorChip, Klokkestua, Norway), hereafter referred to as AmpliCor, HC2 and Proofer, respectively. AmpliCor and HC2 detect the DNA of hrHPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68. Proofer detects and genotypes E6/E7 full-length mRNA transcripts of hrHPV types 16, 18, 31, 33 and 45. For all methods, the result was positive if at least one of the HPV types included therein was detected.

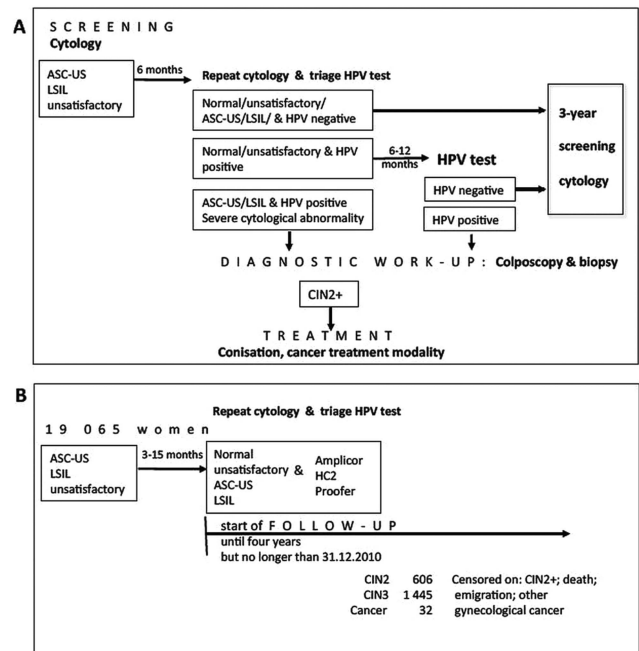


Figure 1 NCCSP recommendations for women with unsatisfactory, ASC-US or LSIL as screening cytology, in 2005–2010 (A) and application of these recommendations to determine follow-up of study participants (B). (A) In 2005–2010, the NCCSP recommended for women with unsatisfactory, ASC-US and LSIL repeat cytology and HPV testing in 6 months. If repeat cytology showed severe abnormalities, or if it showed ASC-US/LSIL in combination with a positive triage-HPV test result, colposcopy and biopsy were recommended. Women with a negative triage-HPV test result combined with ASC-US/LSIL, unsatisfactory or normal repeat cytology were returned to the regular 3-year screening programme. An additional follow-up HPV test was recommended for women with normal or unsatisfactory repeat cytology combined with a positive triage-HPV test result. (B) All 19 065 women who had repeat cytology and HPV test after unsatisfactory, ASC-US and LSIL screening result were divided, based on the triage-HPV method, into mutually exclusive AmpliCor, HC2 and Proofer groups. Individual follow-up started at the date of triage-HPV test and lasted either 4 years, until 31 December 2012, or were censored at time of the histologically confirmed cervical intraepithelial neoplasia grades 2 and 3 (CIN2 and 3), invasive cervical cancer, other gynaecological cancer, death or emigration, whichever came first. AmpliCor, AMPLICOR HPV Test (Roche Diagnostics, Basel, Switzerland); ASC-US, atypical squamous cells of undetermined significance; HC2, Hybrid Capture 2, High-Risk HPV DNA Test (QIAGEN, Gaithersburg, Maryland, USA); HPV, human papillomavirus; LSIL, low-grade squamous intraepithelial lesion; NCCSP, Norwegian Cervical Cancer Screening Programme; Proofer, PreTect HPV-Proofer (NorChip, Klokkestua, Norway).

Use of Norwegian Cancer Registry for identification and follow-up of study subjects

For the present study, in July 2011, relevant information on cervical cancer screening and cervical preinvasive lesions and cancer for 21 958 women were extracted

from databases at the Cancer Registry of Norway. Owing to a 1–2-year delay between diagnosis and complete registration of invasive cancers,¹¹ all cervical cancers detected were verified in early 2012. The personal identification number, an 11-digit numerical code which combines date of birth with five numbers allowing unique identification of every legal resident and citizen in Norway, was used to organise the information on cytology, HPV and histology tests an individual level.

Subjects

Empirical data showed deviation from the NCCSP guidelines and a set of rules was applied to define the combination of cytology and HPV test for clinical decision (figure 2). (1) Per participant a ‘pair’ of repeat cytology and triage-HPV test was identified if a time period between these tests was 2 months or less; (2) pair was used for clinical decision-making if a time period between the triage-HPV test and ASC-US, LSIL or unsatisfactory screening cytology was 3–15 months; (3) when more than one pair could be defined for a participant, we chose the pair closest to the screening cytology; (4) participants with a severe abnormality in repeat cytology were excluded since in this case the triage-HPV test result was irrelevant for referral to colposcopy and biopsy.

In total, 19 065 women were divided into three mutually exclusive HPV groups based on the triage-HPV method used: the Amplicor group (n=4715), the HC2 group (9162) and the Proofer group (5188). Owing to similar performance, in figures 3 and 4, the Amplicor and HC2 groups were combined to a single HPV DNA group.

Statistical analysis

Follow-up time started at the date of triage-HPV test and for all women the worst histologically confirmed lesion was identified: CIN2 (N=606), CIN3 (N=1445), invasive cervical cancer (N=32) or other gynaecological cancer (N=12); death (N=62), emigration (N=6), whichever came first (figure 1B). First smear in follow-up obtained by a non-invasive procedure, that is, cytology or HPV test, and by an invasive procedure, that is, tissue sample, biopsy, cone or endocervical curettage was identified to study adherence to the NCCSP guidelines. All observations were censored at 31 December 2010.

The main results were presented for ‘ASC-US/LSIL’ and ‘normal’ repeat cytology, with the latter including NILM and unsatisfactory repeat cytology. The rationale to combine different cytological diagnostic entities stems from identical management guidelines for these patients and from the study objective, which was to observe the outcome of tests as used in ‘real life’. The online supplementary table includes results separated for ASC-US and LSIL. HPV positivity rate was estimated as proportion of HPV-positives by the HPV group and repeat cytology. We excluded from the analysis 100 inconclusive HPV results, HC2: 4, Proofer: 96 (among

which 11% were diagnosed with the CIN2+). The Kaplan-Meier analysis was used for the cumulative incidence of CIN2+, CIN3+, first non-invasive and first invasive follow-up procedure, among HPV-positive and HPV-negative women at 6 months as short-term follow-up and 3 years as long-term follow-up.

Event history analysis was applied based on the following assumption about censoring: the short-term risk of CIN2+ occurrence during follow-up should remain the same for a given woman, regardless of any previous censoring, which ensures consistency of Kaplan-Meier estimates. This is known as independent censoring in the literature on survival analysis.¹⁴ We also considered the probability of a positive triage-HPV if CIN2+ was diagnosed in the follow-up period before t , that is,

$$S_t = p(\text{HPV} + |\text{CIN2} + \text{detected before } t)$$

To estimate this, we let \hat{X}_t and \hat{Y}_t denote the Kaplan-Meier estimates for the groups with positive and negative triage-HPV, respectively. Moreover, $\hat{\alpha}$ denoted the estimated odds of having a positive test result, regardless of later observations. Bayes’ formula then yields that

$$\hat{S}_t = \frac{1 - \hat{X}_t}{1 - \hat{X}_t + \frac{(1 - \hat{Y}_t)}{\hat{\alpha}}} \approx S_t$$

gives a reasonable estimator for the above triage-HPV detection rate, often referred to as sensitivity. The statistical uncertainty, in terms of 95% CI bands, was calculated by bootstrapping.¹⁵

RESULTS

Following ASC-US, LSIL or unsatisfactory screening cytology, repeat cytology 3–15 months later was normal for 13 673 women, while 5392 had ASC-US/LSIL (table 1). The median follow-up time between screening and triage was 6 months for all groups. Patients with ASC-US/LSIL in the Amplicor group were slightly older, with a median age of 39 years, compared with the HC2 and Proofer groups where IQR of age was 30–46 years, and median age 37 and 38 years, respectively. When the HPV DNA group was compared with the Proofer group, difference in age distribution was not statistically significant ($p=0.665$). Median age of 42 years for the Proofer group in women with normal repeat cytology was slightly higher compared with 41 years in the Amplicor and HC2 groups.

ASC-US/LSIL repeat cytology

Among women with ASC-US/LSIL repeat cytology, 57.3%, 69.9% and 32.6% tested positive by Amplicor, HC2 and Proofer at triage, respectively. Figure 3A shows a rapid increase in the number of biopsies taken shortly after positive triage-HPV DNA or Proofer tests, reaching

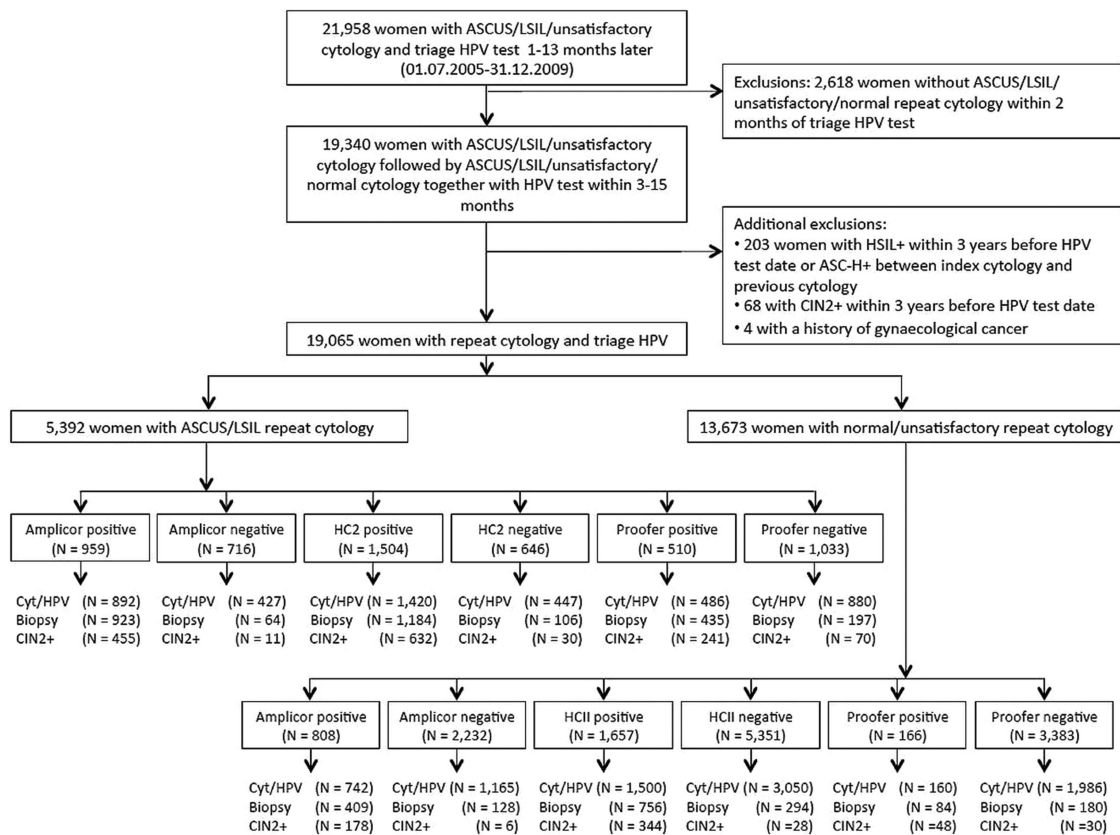


Figure 2 Flow diagram describing selection of and follow-up of study participants (adjusted from proposed Standards for Reporting of Diagnostic Accuracy). From the nationwide registry we identified all 21 958 women 25–69 years of age who had Amplicor, HC2 or Proofer HPV tests taken 1–15 months after an ASC-US, LSIL or unsatisfactory screening cytology during the period of 1 July 2005 to 31 December 2009. While NCCSP recommendations for clinical management, based on HPV and repeat cytology result, are clearly defined, revealed registry linkage a deviation from these recommendations, in terms of timing and frequency of cytology, HPV tests and histology. We used a set of rules to define combination of cytology and HPV test used for clinical decision-making in triage, that is, triage-HPV and repeat cytology: (1) per participant a ‘pair’ of repeat cytology and triage-HPV test was identified if a time period between these test was 2 months or less; (2) pair was used for clinical decision-making if a time period between the triage-HPV test and ASC-US, LSIL or unsatisfactory screening cytology was 3–15 months; (3) when more than one pair could be defined for a participant, we chose the pair closest to the screening cytology; (4) participants with a severe abnormality in repeat cytology were excluded since in this case the triage-HPV test result was irrelevant for referral to colposcopy and biopsy. We defined altogether 19 340 women with the pair of triage-HPV and repeat cytology. Two hundred and seventy-five women were excluded, and 19 065 participants were enrolled in the study. Follow-up time started at the date of triage-HPV test and finished in 31 December 2010. *Endpoint* was the date and diagnosis of the worst histologically confirmed lesion: cervical intraepithelial neoplasia grades 2 and 3 (CIN2 and 3), invasive cervical cancer. To study *adherence to NCCSP guidelines*, we detected (1) non-invasive follow-up event, that is, date of cytology or HPV test, whichever came first, (cyt/HPV) in follow-up and (2) invasive follow-up event, that is, date of first histology diagnosis (biopsy) in follow-up. Follow-up-time was censored at date of other gynaecological cancer, death and emigration, whichever came first. The number of women with non-invasive follow-up event, invasive follow-up event and endpoint is provided by repeat cytology, HPV result and HPV detection method. Those with unsatisfactory HPV results are excluded from this diagram (24 with ASC-US/LSIL and 76 with normal/unsatisfactory). Amplicor, AMPLICOR HPV Test (Roche Diagnostics, Basel, Switzerland); ASC-US, atypical squamous cells of undetermined significance; HC2, Hybrid Capture 2, High-Risk HPV DNA Test (QIAGEN, Gaithersburg, Maryland, USA); HPV, human papillomavirus; LSIL, low-grade squamous intraepithelial lesion; NCCSP, Norwegian Cervical Cancer Screening Programme; Proofer, PreTect HPV-Proofer (NorChip, Klokkarstua, Norway).

about 85% in both groups at 3 years. About 95% of triage-HPV-positives, irrespective of method, had a follow-up cytology/HPV result in 3 years ($p=0.120$) while 64% of triage-HPV DNA-negatives and 88% of triage-Proofer-negatives had a follow-up cytology/HPV result ($p<0.000001$); 12% of triage-HPV DNA-negatives

and 19% of triage-Proofer-negatives had a biopsy ($p<0.00001$; [figure 3E](#)).

Among women with ASC-US/LSIL repeat cytology, 1445 were diagnosed with CIN2+, including 21 cancers ([table 1](#)). The 3-year risk of CIN2+ for positive triage-HPV was 48.1%, 43% and 48.2% in the Amplicor,

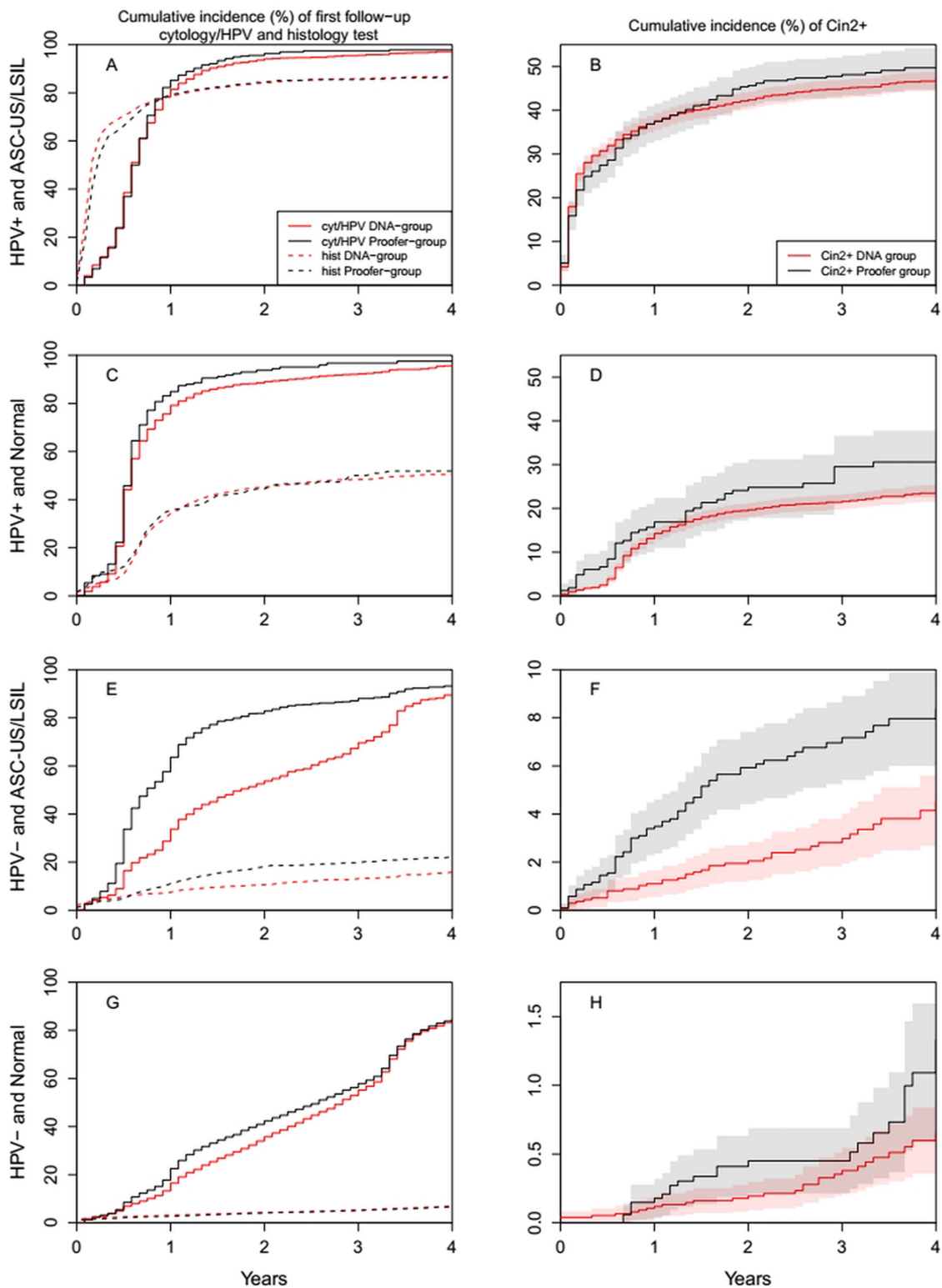
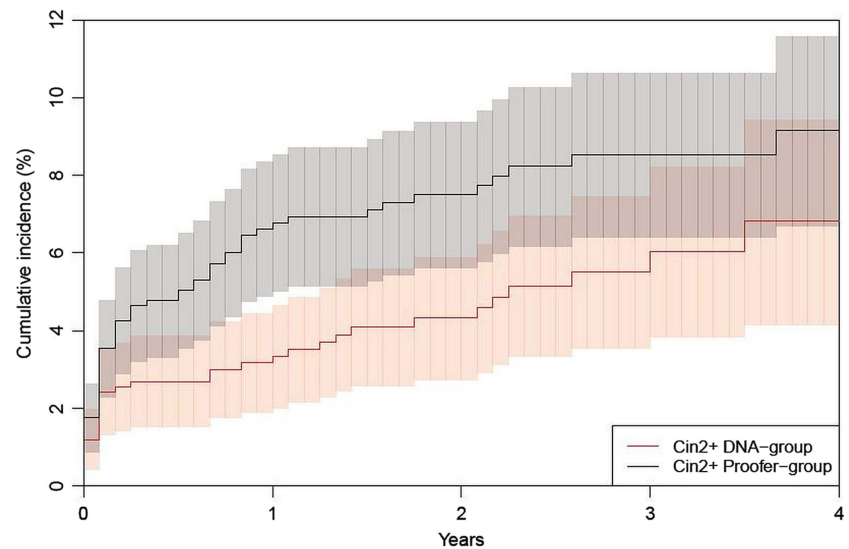


Figure 3 Four-year follow-up of women with ASC-US, LSIL or unsatisfactory screening cytology in 2005–2009 by repeat cytology diagnosis and triage-HPV detection methods. x-axes refer to follow-up time in years from date of triage-HPV; y-axes refer to cumulative incidences as a percentage; red colour refers to HPV DNA and black colour refers to the Proofer group; adherence to non-invasive follow-up (cyt/HPV, solid lines) and invasive follow-up (hist, dotted lines) is depicted in (A), (C), (E) and (G). Four-year risk (cumulative incidence) for CIN2+ with shades denoting 95% CIs is depicted in (B), (D), (F) and (H). (A), (B) and (C), (D) referring to these with positive triage-HPV combined with ASC-US/LSIL and NILM/unsatisfactory in repeat cytology, respectively. (E), (F) and (G), (H) referring to these with negative triage-HPV combined with ASC-US/LSIL and NILM/unsatisfactory in repeat cytology, respectively. ASC-US, atypical squamous cells of undetermined significance; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; HPV, human papillomavirus; LSIL, low-grade squamous intraepithelial lesion; NILM, negative for intraepithelial lesions or malignancy; Proofer, PreTect HPV-Proofer (NorChip, Klokkaerstudia, Norway).

Figure 4 Cumulative incidence of CIN2+ among women who had negative triage-HPV test result combined with ASC-US/LSIL repeat cytology and who had a 'non-recommended' follow-up cytology within 2.5 years. x-Axes refer to follow-up time in years from date of follow-up test; y-axes refer to cumulative incidences; red colour refers to HPV DNA and black colour refers to Proofer group; shades indicate 95% CIs. ASC-US, atypical squamous cells of undetermined significance; CIN2+, cervical intraepithelial neoplasia grade 2+; HPV, human papillomavirus; LSIL, low-grade squamous intraepithelial lesion.



HC2 and Proofer groups, respectively (table 2). There was little difference regarding risk of CIN2+ ($p=0.375$) between HPV DNA-positives and Proofer-positives after 3 years of follow-up (figure 3B). After negative triage-HPV, the 3-year risk of CIN2+ was 2.1%, 4% and 7.2% in the AmpliCor, HC2 and Proofer groups, respectively. Compared with HPV DNA-negatives, the risk among Proofer-negatives was significantly higher for CIN2+ ($p<0.00001$; figure 3F). The risk-ratio for CIN2+ in HPV-positive and HPV-negative women was 22.9 (48.1/2.1) in the AmpliCor group, 10.8 (43/4) in the HC2 group and 6.7 (48.2/7.2) in the Proofer group. Among women with ASC-US/LSIL repeat cytology, negative triage-HPV and follow-up cytology/HPV within 2.5 years, the 2-year risk (after follow-up cytology) for CIN2+ was 4.3% (95% CI 2.7% to 5.7%) for the triage-HPV DNA group and 7.5% (95% CI 5.6% to 9.4%) for the triage-Proofer group ($p=0.023$; figure 4). For all CIN2+ cases diagnosed within 6 months, 99.5%, 97.9% and 90.1% had positive AmpliCor, HC2 and Proofer tests at triage, respectively. For all CIN2+ diagnosed within 3 years, 96.9%, 96.2% and 76.8% were positive for AmpliCor, HC2 and Proofer at triage, respectively (table 2). The triage-HPV detection rate for each HPV method remained largely unchanged when CIN3+ was used as endpoint, and the 3-year risk of CIN3+ after negative triage-AmpliCor, HC2 and Proofer was 1.2%, 3% and 4.4%, respectively.

Normal repeat cytology

Among women with normal repeat cytology, the HPV positivity rate was 26.6%, 23.6% and 4.6% by AmpliCor, HC2 and Proofer, respectively (table 1). The proportion with biopsies was comparable between the two groups for HPV-positive ($p=0.779$) and HPV-negative women ($p=0.581$; figure 3C,G). The 3-year risk of CIN2+ among women with positive triage-HPV was 22%, 21.4% and 29.5% by AmpliCor, HC2 and Proofer, respectively

(table 2). The 3-year risk for severe abnormalities among negative triage-HPV and normal repeat cytology was very low. For all CIN2+ cases diagnosed within 6 months, 97.2%, 95.2% and 100% had positive AmpliCor, HC2 and Proofer tests at triage, respectively. For all CIN2+ cases diagnosed within 3 years, 96.5% and 94.1% had positive AmpliCor and HC2 tests at triage, respectively. Only 76.3% of CIN2+ diagnosed over 3-years were positive by Proofer at triage.

DISCUSSION

In a prospective 3-year follow-up study, we compared the effectiveness of HPV testing with the AmpliCor, HC2 or Proofer, as adjuncts to repeat cytology in the triage of women with ASC-US, LSIL or unsatisfactory screening cytology in the NCCSP. The long-term evaluation of new screening routines showed a good overall performance of triage-HPV DNA testing. However, the management of HPV-negative women with persistent ASC-US/LSIL was suboptimal, as women with residual risk for CIN2+ above the recommended 2% were returned to regular screening. Residual risk was lowest with the HPV-DNA testing. The allocation of HPV detection methods did not depend on the presumed severity of the underlying CIN in this study setting since the Regional Health Trust authorities had predetermined which type of test would be reimbursed. A number of factors may differ between controlled studies and real-world setting, for example, adherence to follow-up recommendations, accuracy of histology and cytology diagnoses as well as HPV testing. Hence, comparative effectiveness research, like the present study, is needed to observe the effect of changes implemented in regular screening.¹⁶ We considered the cumulative incidence of CIN2+ as the most relevant endpoint, as it represents the threshold for treatment, and no further adjudication of the registered histology diagnoses was made.

Table 1 Characteristics of the study population stratified by repeat cytology diagnosis and triage-HPV detection method

Characteristics	Repeat cytology	Amplicor	HC2	Proofer
Number of women				
N	ASC-US/LSIL	1675	2151	1566
N	Normal	3040	7011	3622
Median age				
Years (IQR)	ASC-US/LSIL	39 (32–47)	37 (30–46)	38 (30–46)
Years (IQR)	Normal	41 (33–50)	41 (33–49)	42 (34–51)
No smears in follow-up				
N (%)	ASC-US/LSIL	275 (16.4)	206 (9.6)	141 (9.0)
N (%)	Normal	1057 (34.8)	2296 (32.7)	1294 (35.7)
Number of CIN2+				
N (%)	ASC-US/LSIL	466 (27.8)	661 (30.7)	318 (20.3)
N (%)	Normal	184 (6.1)	373 (5.3)	81 (2.2)
Number of CIN3+				
N (%)	ASC-US/LSIL	307 (18.3)	480 (22.3)	211 (13.5)
N (%)	Normal	120 (3.9)	305 (4.4)	54 (1.5)
Number of cancer				
N (%)	ASC-US/LSIL	6 (0.36)	8 (0.37)	7 (0.45)
N (%)	Normal	2 (0.07)	6 (0.09)	3 (0.08)
HPV positivity rate				
HPV positives (%)	ASC-US/LSIL	959 (57.3)	1504 (69.9)	510 (32.6)
HPV positives (%)	Normal	808 (26.6)	1657 (23.6)	166 (4.6)

Amplicor, AMPLICOR HPV Test (Roche Diagnostics, Basel, Switzerland); ASC-US, atypical squamous cells of undetermined significance; CIN2+, cervical intraepithelial neoplasia grade 2+ (worse); CIN3+, cervical intraepithelial neoplasia grade 3+ (worse); HC2, Hybrid Capture 2, High-Risk HPV DNA Test (QIAGEN, Gaithersburg, Maryland, USA); HPV, human papillomavirus; LSIL, low-grade squamous intraepithelial lesions; normal, including negative for intraepithelial lesions or malignancy and unsatisfactory; Proofer, PreTect HPV-Proofer (NorChip, Klokkestua, Norway); IQR, interquartile range.

Women with ASC-US/LSIL repeat cytology 6 months after ASC-US, LSIL or unsatisfactory screening cytology represent a population with cytological evidence of a persistent HPV infection. The 3-year absolute risk of CIN2+ was high, 43% among HC2-positives and 48% among those positive by Proofer or Amplicor. This is considerably higher than most of the triage studies published to date. For example, among 1578 women with screening-detected ASC-US who underwent colposcopy, an absolute risk of CIN2+ of 14% was reported among hrHPV-positive women, and 31.5% among HPV16 positives,¹⁷ which is a high-risk population as HPV16 is most commonly found in cancers,^{18–20} and because CIN3+ develops sooner following HPV16 detection in baseline than after other HPV genotype detection.^{21–23} Furthermore, a CIN2+ prevalence of 8.7% and 17% in women with screening-detected ASC-US and LSIL cytology, respectively, was reported in a meta-analysis,²⁴ compared with the 29% we report in the HPV DNA group and the 20% in the Proofer group in our study. In contrast, among women who were referred to colposcopy 32.7% harboured CIN2+.²⁵ Observed differences cannot be explained only by different age profiles of study populations, nor variations in the criteria used by different laboratories to diagnose ASC-US and LSIL,²⁶ which is certainly relevant in international comparisons.²⁷ Our population consisting of women with cytological abnormalities persisting at 6 months may be compatible with persistent HPV infection and are, per

se, a strong indicator for CIN2+. The importance of abnormal screening history on CIN2+ risk estimates was also pointed out in a recent publication, where the 4-year risk of CIN2+ among HC2-positive women with cytological abnormalities and a previous abnormal smear was 34.56%.²⁸

Contrary to the national recommendations to return women with ASC-US/LSIL repeat cytology to the regular screening programme if the triage-HPV test is negative, a substantial proportion of HPV-negative women had a follow-up cytology/HPV or histology, with an estimated 3-year risk for CIN2+ of 3.1% in the HPV DNA group and 7.2% in the Proofer group. The observed deviance from the NCCSP recommendations probably reflects scepticism regarding the protective effect of negative triage-HPV in combination with ASC-US/LSIL repeat cytology, as previous guidelines recommended that women with ASC-US/LSIL repeat cytology undergo an additional follow-up cytology in 12 months. Moreover, during the study period (possibly early 2008), the manufacturer of the Proofer test modified their recommendations and suggested follow-up cytology 1 year after a combination of negative triage-Proofer and ASC-US/LSIL repeat cytology,²⁹ which might explain the slightly more intensive follow-up with cytology/HPV of women with negative triage-HPV in the Proofer group compared with the HPV DNA group in the present study. As more intensive biopsy collection leads to higher CIN2+ rates¹⁰ in short

Table 2 Cumulative incidence rates of CIN2+, CIN3+ and triage-HPV detection rate at 6 months and 3 years in women with ASC-US, LSIL or unsatisfactory screening result by repeat cytology in Norway, 2005–2010

	ASC-US/LSIL repeat cytology			Normal/unsatisfactory repeat cytology		
	Amplicor group % (95% CI)	HC2 group % (95% CI)	Proofer group % (95% CI)	Amplicor group % (95% CI)	HC2 group % (95% CI)	Proofer group % (95% CI)
CIN2+						
Cumulative incidence in HPV positives						
6 months	38.8 (35.6 to 41.8)	27.5 (25.3 to 29.8)	28.6 (24.6 to 32.4)	4.3 (2.9 to 5.7)	3.6 (2.7 to 4.5)	8.4 (4.1 to 12.6)
3 years	48.1 (44.7 to 51.3)	43.0 (40.3 to 45.5)	48.2 (43.4 to 52.5)	22.0 (19.0 to 25.0)	21.4 (19.3 to 23.4)	29.5 (21.7 to 36.6)
Cumulative incidence in HPV negatives						
6 months	0.3 (0.0 to 0.7)	1.4 (0.5 to 2.3)	1.5 (0.8 to 2.3)	0.0 (0.0 to 0.1)	0.1 (0.0 to 0.1)	0.0 (0.0 to 0.0)
3 years	2.1 (0.7 to 3.4)	4.0 (2.3 to 5.6)	7.2 (5.4 to 8.9)	0.3 (0.0 to 0.6)	0.4 (0.2 to 0.6)	0.5 (0.2 to 0.7)
Triage-HPV detection rate						
6 months	99.5 (98.9 to 99.8)	97.9 (95.0 to 98.7)	90.1 (82.8 to 93.6)	97.2 (96.5 to 99.3)	95.2 (85.5 to 98.0)	100 (100 to 100)
3 years	96.9 (92.6 to 98.1)	96.2 (94.0 to 97.5)	76.8 (70.7 to 81.6)	96.5 (88.9 to 98.3)	94.1 (90.0 to 96.1)	76.3 (60.6 to 84.8)
CIN3+						
Cumulative incidence in HPV positives						
6 months	25.4 (22.6 to 28.2)	18.9 (16.9 to 20.9)	18.2 (14.8 to 21.5)	2.2 (1.2 to 3.2)	2.5 (1.7 to 3.2)	4.2 (1.1 to 7.2)
3 years	31.8 (28.7 to 34.8)	31.2 (28.7 to 33.6)	32.3 (27.9 to 36.4)	14.7 (12.1 to 17.3)	17.7 (15.7 to 19.6)	21.9 (14.8 to 28.4)
Cumulative incidence in HPV negatives						
6 months	0.0 (0.0 to 0.0)	0.5 (0.0 to 1.0)	1.1 (0.4 to 1.7)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)
3 years	1.2 (0.1 to 2.2)	3.0 (1.6 to 4.5)	4.4 (3.0 to 5.7)	0.2 (0.0 to 0.4)	0.3 (0.1 to 0.5)	0.3 (0.1 to 0.4)
Triage-HPV detection rate						
6 months	100 (100 to 100)	99.0 (96.8 to 99.6)	89.4 (79.5 to 93.7)	100 (100 to 100)	100 (100 to 100)	100 (100 to 100)
3 years	97.4 (89.8 to 98.7)	96.0 (92.5 to 97.2)	78.5 (71.2 to 83.9)	96.8 (87.9 to 98.6)	95.0 (88.9 to 97.0)	81.0 (59.1 to 89.7)

Amplicor, AMPLICOR HPV Test (Roche Diagnostics, Basel, Switzerland); ASC-US, atypical squamous cells of undetermined significance; CIN2+, cervical intraepithelial neoplasia grade 2+ (worse); CIN3+, cervical intraepithelial neoplasia grade 3+ (worse); HC2, Hybrid Capture 2, High-Risk HPV DNA Test (QIAGEN, Gaithersburg, Maryland, USA); HPV, human papillomavirus; LSIL, low-grade squamous intraepithelial lesion; Proofer, PreTect HPV-Proofer (NorChip, Klokkarstua, Norway).

follow-up studies, it is possible that the more intensive follow-up observed in the Proofer group partially explains the differences in the risk of CIN2+ among triage-Proofer-negative and triage-HPV DNA-negative women. However, our subanalysis, restricted to women with negative triage-HPV who had follow-up cytology/HPV (for any reason, figure 4), also found a higher risk for CIN2+ in the Proofer group. In addition, we observed higher rates of CIN2+ in the Proofer group compared with the HPV DNA group among women with normal repeat cytology, regardless of intensity of biopsy collection after follow-up cytology/HPV. Thus, the differences in CIN2+ incidence rates among women testing negative for the different tests cannot be entirely explained by differential ascertainment. One can argue that CIN3+ as endpoint should be preferred for CIN2+ as higher CIN2+ rates may reflect mostly self-limiting lesions. We observed lower 3-year risk for CIN3+ compared with CIN2+ in HPV-negative ASC-US/LSIL for all HPV methods but the difference between the AmpliCor and the Proofer remained, which would suggest inferior performance of the Proofer also with CIN3+.

A risk for CIN2+ below 2% has been proposed as a cut-off for safe return to regular screening programmes in a risk model for clinical decision-making, launched recently to standardise care for maximal patient safety.³⁰ Hence, our results suggest that only women with negative triage-AmpliCor had a CIN3+ risk low enough to justify their return to the regular screening programme. It might be that HPV test performance can be improved if careful monitoring of separate labs actually performing an HPV test is implemented. In contrast, HPV-negativity among women with normal cytology after ASC-US, LSIL or unsatisfactory screening cytology predicted a very low CIN2+ risk in 3 years (<1%). This is in line with prospective natural history studies, which showed a low risk of CIN3+ in hrHPV DNA-negative women,^{23 31} and several studies evaluating HPV testing in primary screening.^{32 33}

The importance of long-term follow-up of real-world implementation of new technologies when estimating the accuracy of a test in a cancer prevention programme has been pointed out repeatedly.^{34 35} Screening recommendations in a particular programme, and most importantly, how these recommendations are actually carried out, define the needed length for follow-up studies with CIN2+ as an endpoint. This is well illustrated by our data where the cumulative incidence curve for first biopsy and CIN2+ are similar in shape. Traditional indicators of test properties, such as sensitivity, specificity, negative and positive predictive value, cannot account for variations in the follow-up time of each individual. To overcome these methodological restrictions, we applied cumulative incidences to assess how risk for severe abnormalities varied over time, given set combinations of test results, and used it to estimate the proportion of HPV-positives at baseline for CIN2+ patients diagnosed during the course of follow-up as an

approximation of (longitudinal) sensitivity. Considering all CIN2+ diagnosed during the 3-year follow-up, 97% were positive for hrHPV DNA when repeat cytology diagnosis was ASC-US/LSIL, and 95% were positive when repeat cytology was normal. This is in line with the reported sensitivity for CIN2+ of 93.1% for HC2 and 97.2% for AmpliCor in ASC-US and LSIL triage.^{25 36} The probability that a patient diagnosed with CIN2+ had a positive triage-HPV test result was higher among cases diagnosed within 6 months than in 3 years, irrespective of HPV detection method. However, after 3 years, we observed an obvious drop in the proportion of CIN2+/CIN3+ patients positive for HPV at triage in the Proofer group, indicating that in the NCCSP, hrHPV DNA detection methods had a better prognostic value over time for CIN2+ than the method detecting only the mRNA of five hrHPV types. HPV16, 18, 31, 33 and 45 are reported to be present in 75.1% of CIN2/3 and 88.5% of squamous cell carcinomas.^{19 37 38} This fits well with our estimated triage-Proofer detection rate for CIN2+ of 76%. Fewer women were referred to biopsy when the Proofer was used, compared with hrHPV DNA detection methods, reducing the risk of side effects of conisation. However, the residual risk of not detecting CIN2+ when using the Proofer was substantial.

In the current study, we provide risk estimates for CIN2+ to address the need for real-life evaluation of management strategies for patients with a given set of test results. The Proofer had lower estimated sensitivity for CIN2+ in triage compared with hrHPV DNA detection methods. When we used cancer as the end point, the estimates were highly imprecise due to low numbers of observations, but resembled the estimates seen with the CIN2+ and CIN3+ endpoints (data not shown).

Our findings have several important implications not only for the NCCSP, but also for other countries in Europe which follow similar guidelines and use same HPV detections methods, albeit the Proofer may mostly be used in Scandinavia.^{39 40} Evaluation of guidelines, which currently recommend that women with ASC-US, LSIL or unsatisfactory screening cytology undergo repeat cytology and HPV-testing in 6 months showed that

1. Triage-HPV DNA had a high sensitivity for CIN2+ among women with ASC-US/LSIL repeat cytology. However, residual risk for CIN2+ in repeat ASC-US/LSIL and triage-HPV-negatives was over 2%. Return of these women to the normal screening programme is potentially unsafe, which underlines the need for improved and more specific guidelines in Norway in management of women with ASC-US or LSIL repeat cytology.
2. Negative triage-HPV in combination with normal repeat cytology resulted in a low-residual risk of severe abnormalities before the next screening round. However, the triage-HPV positivity rate among CIN2+ and CIN3+ was higher when detecting hrHPV DNA than mRNA from five hrHPV types, suggesting the priority of methodology detecting all hrHPV.

Generally, follow-up time needed for studies evaluating screening programme performance is dependent on the programme's particular recommendations, and studies with a short follow-up time cannot provide sufficient information about screening programme performance. Comparative effectiveness research is essential to study the effect of HPV testing, its benefits and harms, in a real-world setting.^{41 42} As our study exemplifies, screening programmes should, when implementing new policies, be prepared to continuously evaluate performance using nationwide and comprehensive comparative effectiveness research, and to modify their policies accordingly to provide the best healthcare possible.

Acknowledgements The authors would like to thank Anna Skog for technical support and critical comments during preparation of this manuscript.

Contributors MN, KR and JD contributed substantially to conception and design and interpretations of data. SC contributed substantially to analysis and interpretation of data. Anna Skog contributed to drafting the article and data analysis. MN, KR, JD and SC contributed to drafting the article and revising it critically for important intellectual content and approved the final version.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All authors, external and internal, had full access to all of the data in the study (including statistical reports and tables) and can take responsibility for the integrity of the data and the accuracy of the data analysis. Anonymous individual-level data and full dataset are available at Cancer Registry of Norway. Please contact the Data Enquiry unit (Datautleverings enhet, DUE) for further details at datautlevering@krefregisteret.no

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

1. 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.
2. 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.
3. zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* 2002;2:342–50.
4. Rijklaart 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.
5. 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;62:147–72.
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. Johansen BK, Skare G, Haldorsen T, *et al.* *Masseundersøkelsen mot livmorhalskreft. Årsrapport 2008: Kreftregisteret.* Institutt for populasjonsbasert kreftforskning, 2009.
8. 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.
9. Nygard JF, Skare GB, Thoresen SO. The cervical cancer screening programme in Norway, 1992–2000: changes in Pap smear coverage and incidence of cervical cancer. *J Med Screen* 2002;9:86–91.
10. Nygard JF, Nygard M, Skare GB, *et al.* Pap smear screening in women under 30 in the Norwegian Coordinated Cervical Cancer Screening Program, with a comparison of immediate biopsy vs Pap smear triage of moderate dysplasia. *Acta Cytol* 2006;50:295–302.
11. Larsen IK, Smastuen M, Johannesen TB, *et al.* Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *Eur J Cancer* 2009;45:1218–31.
12. Solomon D, Nayar R. eds. *The Bethesda System for reporting cervical cytology.* 2nd edn. New-York: Springer-Verlag New-York, Inc., 2004.
13. Tavassoli F, Devilee P. eds. *Pathology and genetics of tumours of the breast and female genital organs.* Lyon: IARC Press, 2003.
14. Andersen PK, Borgan O, Gill R, *et al.* *Statistical models based on counting processes.* New York: Springer-Verlag, 1993.
15. Boot: Bootstrap R (S-Plus) Functions [program]. R package version 1.3–4 version, 2012.
16. Sullivan P, Goldmann D. The promise of comparative effectiveness research. *JAMA* 2011;305:400–1.
17. 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.
18. de Sanjose S, Quint WG, Alemany L, *et al.* Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 2010;11:1048–56.
19. Smith JS, Lindsay L, Hoots B, *et al.* Human papillomavirus type distribution in invasive cervical cancer and high-grade cervical lesions: a meta-analysis update. *Int J Cancer* 2007;121:621–32.
20. Wheeler CM, Hunt WC, Joste NE, *et al.* Human papillomavirus genotype distributions: implications for vaccination and cancer screening in the United States. *J Natl Cancer Inst* 2009;101:475–87.
21. 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.
22. Castle PE, Schiffman M, Wheeler CM, *et al.* Human papillomavirus genotypes in cervical intraepithelial neoplasia grade 3. *Cancer Epidemiol Biomarkers Prev* 2010;19:1675–81.
23. Kjaer SK, Frederiksen K, Munk C, *et al.* Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: role of persistence. *J Natl Cancer Inst* 2010;102:1478–88.
24. Cuzick J, Arbyn M, Sankaranarayanan R, *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.
25. Szarewski A, Mesher D, Cadman 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.
26. Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA* 2001;285:1500–5.
27. Scott DR, Hagmar B, Maddox P, *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.
28. Castle PE, Fetterman B, Poitras N, *et al.* Variable risk of cervical precancer and cancer after a human papillomavirus-positive test. *Obstet Gynecol* 2011;117:650–6.
29. Lie AK, Lømo J, Sauer T, *et al.* HPV DNA versus mRNA testing in triage among Norwegian women. In: 27th_IPV_Conference. *Poster presentation.* Berlin, 2011.
30. Castle PE, Sideri M, Jeronimo J, *et al.* Risk assessment to guide the prevention of cervical cancer. *Am J Obstet Gynecol* 2007;197:356 e1–6.
31. 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.
32. Bulkman NW, Berkhof J, Rozendaal L, *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;370:1764–72.
33. Nauciler 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;101:88–99.

34. Cuzick J. Long-term follow-up in cancer prevention trials (It ain't over 'til it's over). *Cancer Prev Res (Phila)* 2010;3:689–91.
35. Wahlstrom C, Iftner T, Dillner J, *et al*. Population-based study of screening test performance indices of three human papillomavirus DNA tests. *J Med Virol* 2007;79:1169–75.
36. 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.
37. Insinga RP, Liaw KL, Johnson LG, *et al*. A systematic review of the prevalence and attribution of human papillomavirus types among cervical, vaginal, and vulvar precancers and cancers in the United States. *Cancer Epidemiol Biomarkers Prev* 2008;17:1611–22.
38. Munoz N, Bosch FX, de Sanjose S, *et al*. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 2003;348:518–27.
39. *Screening for livmoderhalskræft-anbefalinger*. København: Sundhedsstyrelsen, 2012. ISBN:978-87-7104-210-8
40. Lönnberg S, Skare G. *Masseundersøkelsen mot livmorhalskreft. Årsrapport 2009–2011*. Oslo: Kreftregisteret. Institutt for populasjonsbasert kreftforskning, 2013.
41. Iglehart JK. Prioritizing comparative-effectiveness research—IOM recommendations. *N Engl J Med* 2009;361:325–8.
42. Hakama M, Malila N, Dillner J. Randomised health services studies. *Int J Cancer* 2012;131:2898–902.