Methylation in predicting progression of untreated high-grade cervical intraepithelial neoplasia

Karolina Louvanto, PhD[1,2,3†], Karoliina Aro, MD[3†], Belinda Nedjai, PhD [2], Ralf Bützow, PhD [4], Maija Jakobsson, PhD [3], Ilkka Kalliala, PhD[3,5], Joakim Dillner, PhD [6], Pekka Nieminen, PhD [3††], Attila Lorincz, PhD [2††]

[1] Department of Obstetrics and Gynecology, Turku University Hospital, University of Turku, Turku, Finland.

[2] Center for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, UK.

[3] Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.

[4] Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.

[5] Department of Surgery & Cancer, IRDB, Faculty of Medicine, Imperial College, London, UK.

[6] Department of Laboratory Medicine, Karolinska Institute, Stockholm, Sweden.

[†] Authors Karolina Louvanto and Karoliina Aro contributed equally to this manuscript

[††] Authors Pekka Nieminen and Attila Lorincz contributed equally to this manuscript

Corresponding author: Professor Attila Lorincz; Center for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK, email: a.lorincz@qmul.ac.uk, tel: +44 (0) 20 7882 3540

Summary: Many cervical intraepithelial neoplasias (CIN) regress spontaneously or persist but to date there is no method to predict progression. In this study we show that a DNA methylation classifier can predict progression or regression of untreated CIN grade 2.

Abstract

Background

There is no baseline prognostic test to ascertain whether cervical intraepithelial neoplasia (CIN) will regress or progress. The majority of CIN regress in young women and since local treatments are known to increase the risk of adverse pregnancy outcomes interventions need to be sparing. We investigated the ability of a DNA methylation panel (the S5-classifier) to discriminate between progression and regression among women of childbearing age with untreated CIN grade 2 (CIN2).

Methods

Pyrosequencing methylation and HPV genotyping assays were performed on exfoliated cervical cells from 149 young women with CIN2 in a 2-year cohort study of active surveillance.

Results

Twenty-five lesions progressed to CIN grade 3 or worse, 88 regressed to less than CIN grade 1, and 36 lesions persisted as CIN1/2. When cytology, HPV16/18- and HPV16/18/31/33-genotyping, and S5 at baseline were compared to outcomes, S5 was the strongest biomarker associated with regression versus progression. S5 alone or in combination with HPV16/18/31/33-genotyping also showed significantly increased sensitivity versus cytology, comparing regression vs. persistence/progression. With both S5 and cytology tests set at a specificity of 38.6% (95% CI 28.4-49.6) the sensitivity of S5 was significantly higher (83.6%, 95% CI 71.9-91.8) than for cytology (62.3%, 95% CI 49.0-74.4) (p=0.005). The highest area under the curve (AUC) was 0.735 (95% CI 0.621-0.849) in the regression vs.

progression outcome with a combination of S5 and cytology, whereas HPV16/18 or HPV16/18/31/33-genotyping did not provide additional prognostic information.

Conclusions

The S5-classifier shows high potential as a prognostic biomarker to identify women with progressive CIN2.

Key words:

DNA methylation, cervical intraepithelial neoplasia, CIN, high grade squamous intraepithelial lesion, HSIL

Introduction

Cervical intraepithelial neoplasia (CIN) is caused by persistent HPV-infection which is common in women in reproductive age. Most HPV-infections and even CIN can regress without treatment.[1,2] The mildest low grade squamous intraepithelial lesion(LSIL, formerly CIN grade 1,CIN1) are treated with expectant management for up to two years before proceeding to local treatments in persistent disease.[3] Some guidelines also suggest this approach for CIN grade 2(CIN2) in young women.[3] Recent evidence shows that 60% of CIN2 regress spontaneously within two years while only 11% progress in women under the age of 30.[4] Overtreatment of any lesions should be avoided, especially in young women, as treatment significantly increases the risk of adverse outcomes in subsequent pregnancies.[5–8]

To date there is no prognostic test to ascertain whether a CIN lesion has a tendency to regress or progress, leaving treatment algorithms dependent on repeated examinations and testing. The HPV-genotype does not appear to have enough predictive potential on the outcome. Although increasing proportions of severe lesions are caused by HPV16/18 the progressive potential remains mostly uncertain.[9,10] Immunostaining of histological samples, for example with p16, show increasing positivity with increasing severity of lesions, but have not been found consistently prognostic.[11–13]

DNA methylation of both HPV and host genes have been shown to increase with increasing severity of lesions.[14–18] Methylation as a screening triage to high-risk HPV(hrHPV) positive women has also been found to be promising in predicting high-grade CIN(CIN2/3+).[17,19–21] The usefulness of DNA-methylation status in predicting the

outcomes of prevalent cervical lesions has not been shown in a prospective longitudinal series. Here we present results of the prognostic potential of a DNA-methylation biomarker panel in a prospective cohort study of expectant management of untreated histologically confirmed CIN2 in young women.

Methods

Patients and study protocol

The study is part of an ongoing prospective cohort at the Colposcopy Unit of Helsinki University Hospital, Finland. Eligible women diagnosed with histological CIN2 were given written information on active surveillance as an alternative to loop electrosurgical excision procedure(LEEP). The protocol, inclusion and exclusion criteria of the study are shown in Figure 1.

At the baseline visit, after written consent, LEEP was not performed, but instead a new colposcopy was done to ensure eligibility. A cervical brush sample was obtained for HPV-genotyping and methylation-analyses. Scheduled study visits were at 6,12,18, and 24-months. LEEP was performed if histopathological progression to CIN3 or cancer(CIN3+) was observed at any visit or if persistence (CIN2 or CIN1) was observed at the 24-month visit. LEEP was also performed on patient request or if the patient moved out of the region. All histopathological and cytological samples were reviewed by the institution's pathologists and a second opinion was provided on all baseline biopsies by an expert pathologist(RB). If the histopathological diagnosis was not agreed upon the patient was excluded from the study and treated according to Finnish treatment guidelines.[22]

The study protocol was approved by Helsinki University Hospital's Ethical Committee(131/13/03/03/2013;24/4/2013) and was registered in the ISRCTN-registry(ISRCTN91953024). Our study was conducted following REMARK-guidelines.[23]

Sample processing and HPV-genotyping

At the baseline visit, the cells collected in Sample Transport Medium(STM) (Qiagen GMBH, Germany) were stored at −20°C. The samples were later divided into three aliquots and stored at −80°C. HPV-genotyping was done at the Karolinska Institute, Stockholm, Sweden, with the Luminex-assay as previously described.[24]

Methylation analyses

DNA was extracted from aliquots of the STM with the QIAamp DNA Mini Kit(Qiagen Inc, Hilden, Germany). Two hundred nanograms of DNA were used in the bisulfite conversion reactions, where unmethylated cytosines were converted to uracil with the EZ DNA-methylation kit(Zymo research, Irvine,USA). Converted DNA from an equivalent of 1600cells per sample were amplified by methylation- independent PCR-primers and the amplicons were tested by pyrosequencing for DNA-methylation of *EPB41L3* and the late(L1 and/or L2) regions of HPV16,HPV18,HPV31 and HPV33, as previously described.[19,25,26] The laboratory was blinded to HPV-genotyping results; therefore, each methylation-assay in the S5-classifier was run on all specimens. No HPV-type inconsistencies were detected on the methylation versus the HPV-genotyping results.

Statistical analyses

The primary clinical contrast was progression to CIN3+ versus regression to <CIN1. We also evaluated the former two versus persistence (CIN2 or conversion to persistent CIN1). Only women whose histopathological diagnoses changed to <CIN1 and remained constant during all their follow-up visits where considered as regressed. If a woman's normal histology subsequently progressed back to CIN1 or CIN2 she was categorized as persistent for CIN. Our primary question was whether the S5-classifier or the different methylation biomarkers within the classifier could predict progression to CIN3+ among women with CIN2. The clinical outcome groups were defined according to the histopathological findings during follow-up. The S5-classifier was defined as S5=30.9(EPB41L3)+13.7(HPV16L1)+4.3(HPV16L2)+8.4(HPV18L2)+22.4(HPV31L1)+20.3(HPV33L2) with individual CpG-sites as described previously.[21,27] DNA-methylation status for baseline missing values (n=8) of HPV16 were imputed with the value of zero for any HPV16 negative sample (n=5) and by the median for HPV16-positive samples(n=3). Missing values for *EPB41L3* (n=8) were imputed by the median independently of their HPV-infection status. A total of 8 women with missing HPV-genotyping results were imputed as HPV-negative. Our main measures of performance were odds ratios(OR), sensitivity and specificity comparisons, with cut-offs for S5 set at the upper tertile(upper 1/3 of methylation-levels) or at the predefined and validated cut-point of S5=0.8. Cytology was categorized according to the Bethesda-classification. HPV16,HPV18,HPV31 and HPV33-combinations were regarded as binary-variables, with any of these types detected regarded as a positive result versus an all HPV-negative result.[19,27]

Differences between baseline characteristics and mean methylation levels in the three

different clinical outcome groups were compared with Mann-Whitney or Fisher's exact-test or nonparametric-test for trend, as applicable. The Cuzick-test for trend was used to compare the mean methylation levels of the different markers among the diagnostic groups. Unconditional logistic regression OR and 95% confidence interval(CI) were used to evaluate the associations of mean methylation level or the upper tertile-level of different methylation markers and various clinical outcome comparisons. The high tertile of methylation was defined as any value within the upper 1/3 of the distribution of methylation values identified for each methylation biomarker in the specific outcome category comparison.[14] Multivariable models of logistic regression were used to evaluate possible confounding-factors in the methylation versus clinical outcome comparisons and to investigate different biomarker associations between the various clinical outcome groups.

Difference in sensitivity at a selected methylation test cut-point, where the specificity of the methylation test was held equal to the reference comparator (either cytology, HPV16/18-genotyping or both) was assessed by McNemar's-test. The performance of different methylation markers and screening protocols was measured by receiver operator characteristic(ROC)-analysis, by comparing area under the curve(AUC). Kaplan-Meier curves were used to assess the cumulative proportions of women who progressed to CIN3+ by time(in months) since the diagnosis of CIN2. In this analysis persistent CIN1/2 was regarded as non-progression. A likelihood-ratio test was used to assess differences between women with all positive biomarkers to women who tested negative for all markers. Cox proportional hazards regression models were used to estimate unadjusted Hazard Ratios(HR) (95%CI) to examine associations between median methylation and CIN2 progression (date of CIN3+ diagnosis). All p-values were two sided and $p \leq 0.05$ was regarded as significant. Statistical analyses were performed using Stata15(STATA Corp., College Station,TX).

Results

A total of 149 women with histologically confirmed CIN2 and at least two (six-monthly) follow-up visits were included. A flow chart of the study is shown in Figure 1. From the 149 women, 147 women had a follow-up visit at six months (two women had the first follow-up visit at month 12). 116 women had two follow-up visits (at 6 and 12-months). 52 women had an 18-month visit, and 65 women completed the full schedule of follow-up visits to 24-months. All 25 women (17%) who progressed to CIN3+ were treated by LEEP. Of 88 women (59%) categorized as regressed to <CIN1 42 exited the study without treatment and the remaining women are still under follow-up. Of the 36 women (24%) categorized as persistent with CIN1/2, seven LEEP-procedures were performed at the end of the 24-month period, and all had histological CIN2. The remaining 12 women in the persistent group at the 24-month visit had CIN1 and are being followed up according to clinical guidelines. The other remaining 17 women in the persistent group have not yet completed all the follow-up visits.

Baseline characteristics of the women are presented in Table 1. The mean age was 26 years (range: 25.9 to 27.0 years) and did not differ significantly between the three outcome groups. Twenty-one of the 25 women who progressed to CIN3 were hrHPV(one or more of types: 16,18,31,33,35,39,45,51,52,56,58,59)-positive. In contrast, 63 of the 88 women who regressed to <CIN1 were hrHPV-positive. Finally, 32 of the 36 women who persisted as CIN1/2 were hrHPV-positive. Overall 82.3% (116/141) of the women were positive for hrHPV, and of these 52.6%(61/116) were positive for HPV16, while 9.5%(11/116) were positive for HPV18. There was a significant difference(p=0.02) between the regression and persistence group considering hrHPV-positivity, 94.1% of women who persisted were

positive compared to 75.0% in the regression group.

Mean methylation levels of the host gene *EPB41L3* (CpG-sites 438,427,425), the viral HPV16L1-gene (CpG-sites 6367,6389), and the S5-classifier according to clinical outcome are presented in Supplementary Figure 1. Statistical significance in pairwise comparisons of progression to CIN3+ versus regression to <CIN1 was found with *EPB41L3* alone(p=0.02), while for the full S5-classifier the difference was highly significant(p=0.001).

Table 2 presents results for clinical outcome comparisons between each individual outcome (regression, progression, persistence, and combinations of the latter with the former two) for S5, *EPB41L3* and the HPV16L1-methylation biomarkers, using either the pre-validated high tertile levels or the median methylation as cut-offs.[14] For the high tertile cut-off S5 reached statistical significance in almost all comparisons(except persistence vs. progression), with the highest OR of 4.84 (95%CI 1.35-17.41) and an AUC of 0.718 (95%CI 0.61-0.83) observed for regression vs. progression. ORs for mean methylation were found to be significant in all clinical outcome comparisons except regression vs. persistence for the S5-classifier and for *EPB41L3* alone, except for the intermediate group comparisons of persistence and regression or progression. HPV16L1-methylation alone did not show a significant association with any of the comparisons.

We explored the performance of a previously validated high tertile cut-off for the S5-classifier and compared this variable to a cytology cut-point of ≤ASC-US vs. ≥LSIL, and to HPV16/18-and HPV16/18/31/33-positive versus negative (with regression as the referent group;Table 3). The S5-classifier showed the highest significant association with progression OR of 3.39(95%CI 1.35-8.50) followed by HPV16/18/31/33-genotyping with an OR of

3.17(95%CI 1.15-8.68). HPV16/18/31/33-genotyping had also a significant association to CIN2 persistence which the other markers did not show giving an OR of 3.50(95%CI 1.44-8.52).

In a multivariable model comparing OR's of *EPB41L3* and S5-classifier to different clinical outcomes, only S5-classifier showed to be independent predictors of outcomes among the regression vs. persistence and regression/persistence vs. progression groups (adjusted for HPV16/18/31/33-status, abnormal cytology, smoking status, and age) (Supplementary Table 1).

Sensitivity, specificity, positive and negative predictive values were compared between reference tests (cytology at varying cut-points and HPV16/18-positivity or negativity) and the index methylation marker *EPB41L3* alone and the S5-classifier in the outcome of regression vs. persistence/progression(Supplementary Table 2). Our comparisons focused on sensitivity differences when the index test cut-offs were set to allow the closest approximation of specificity between the index and the comparator reference test. The S5-classifier showed significantly increased sensitivity compared to cytology (cut-point less than or equal to atypical squamous cells of undetermined significance(≤ASC-US) vs. ≥LSIL) with a sensitivity of 86.9% (95%CI 75.8-94.2) for S5 versus 75.4% (95%CI 62.7-85.5) for cytology ≥LSIL(p=0.05). In contrast, a cytology cut-point of negative for intraepithelial lesion or malignancy(NILM) vs. ≥ASC-US was essentially non-specific for progression thereby producing an unrealistic comparison to S5. With a cytology cut-point at ≤LSIL vs. high grade squamous intraepithelial lesion or worse(≥HSIL) and a set specificity of 38.6%(95%CI 28.4-49.6) the sensitivities were significantly different at 83.6%(95%CI 71.9-91.8) for S5 and 62.3%(95%CI 49.0-74.4) for cytology(p=0.005). HPV16/18-genotyping performed similarly

to the S5-classifier when specificity of both were made equal(p=1.00). However, this comparison produced a maximum sensitivity of 57%, which we regard as too low for a prognostic biomarker. Indeed, the S5-classifier could be set to a much higher sensitivity (>75%) albeit with a commensurate loss in specificity compared to HPV16/18-genotyping(Figure 2). HPV16L1-methylation and HPV16-genotyping markers were also tested but showed a much lower diagnostic utility compared to the *EPB41L3* and S5-classifier among the clinical outcome comparison of regression vs. persistence/progression.

The performance of the S5-classifier alone and in combination with other tests(≥HSIL cytology and/or HPV16/18-or 16/18/31/33-positivity) was tested in different clinical outcome categories either separately or grouped(Figure 2). The highest AUC was 0.735(95%CI 0.621-0.849) in the regression vs. progression clinical outcome comparison with a combination of S5 above a cut-off of 0.8 and cytology ≥HSIL regarded as positive. Combining HPV16/18 or HPV16/18/31/33-positives with S5 and cytology did not provide any additional advantage. This was seen in all the clinical outcome comparison groups, expect with regression vs. persistence/progression group where combining HPV16/18/31/33 with S5 gave the highest AUC of 0.666 (95%CI 0.580-0.752). Comparisons of *EPB41L3*, S5 and HPV16-positive in clinical outcome comparisons of persistence vs. progression and regression vs. persistence/progression are presented in Supplementary Figure 2. It is noteworthy that the S5-classifier alone provided better performance in discriminating the clinical outcome of progression vs regression, whereas HPV16/18/31/33-positivity performed better in predicting persistent HPV-infection.

Figure 3 shows a significant difference(LR-test p=0.03) between cumulative proportions of progression to CIN3+ distributed by time in women positive for S5, HPV16/18, and cytology

≥HSIL vs. negative for all the previous tests (S5-classifier ≤0.8, HPV16/18-negative, and cytology<HSIL). Cox proportional hazards regression modelling was further used to estimate the HR to examine associations between median methylation of S5-classifier and CIN2 progression (date of the CIN3+ diagnosis) among the group of regression/persistence vs. progression. The HR for S5-classifier alone was 4.19(95%CI 1.57-11.17) and HR was 3.84(95%CI 1.13-13.04) when adjusted with ≥HSIL cytology, HPV16/18-positivity, age, and smoking status.

Discussion

This is the first study to assess the predictive potential of the S5 DNA-methylation classifier in a prospective longitudinal series of patients with histological CIN2 at baseline. We found the S5-classifier to be a significant predictor of progression versus regression in women with untreated CIN2, even after adjusting for cytology, HPV16/18/31/33-genotyping, age, and cigarette smoking. *EPB41L3* and HPV16 L1 individually did not perform similarly with the former biomarker being much better than the latter for progression as the main outcome. The sensitivity of S5 was significantly higher than cytology with varying cut-offs (≤ASC-US vs. ≥LSIL and ≤LSIL vs. ≥HSIL) in assessment of clinical outcomes regression vs. persistence/progression. With a high tertile cut-off value for methylation the OR's in favor of S5 prognostic potential were even higher and the greatest AUC 0.735 (95%CI 0.621-0.849) was achieved when S5 was combined with ≥HSIL cytology.

Although HPV16/18/31/33-genotyping was as good as S5 in predicting regression versus the combination of persistence/progression (Figure 2C) it was not as good as S5 in predicting progression versus regression (Figure 2A and Table 3). The equivalence of S5 to

HPV16/18/31/33 prediction for the combination of persistence and progression categories appears to be driven mainly by the relatively larger persistence group. It should be considered that the natural history of long-term HPV persistence with respect to eventual true progression of CIN2 to CIN3+ versus regression to normal beyond 2 years remains unclear. In our clinical setting persistence of CIN1/2 for two years was taken as an indication for treatment, however, we don't know what proportion of these treatments were really necessary to prevent cervical cancer.

We have a unique study population and CIN2 management strategy. The strengths of our study include focus on a very important clinical question, ability to conduct vigilant follow-up by an expert medical team and expert histopathological diagnosis of CIN-categories. We also have a careful parallel comparison of the different methylation panels and other comparison tests to minimize bias. A weakness is that our results cannot be directly generalized to other histopathological diagnoses. Also, our study was restricted to young women and the length of follow-up varied, which may result in some reclassification of outcomes(regression, persistence) as follow-up continues. This was especially the case regarding the persistent category as the true nature of their disease remains undefined.

The S5-classifier has previously been proven to identify risk of ≥CIN2/3 in hrHPV-positive women in cross-sectional studies based on screening and colposcopy populations where the classifier outperformed triage with HPV16/18-genotyping.[19,27] In the current study we show that S5 can differentiate between regressive and progressive CIN2. Methylation of other combinations of host and HPV-genes have also been found to increase proportional to with severity of lesions,[17] but this has not been examined in a longitudinal series of patients with CIN except a small series of HIV-positive women, where patients with

persistent CIN2/3 had higher methylation of *EPB41L3* than women with <CIN1 or regression to <CIN1.[28] Another host gene FAM19A4 has been shown to be more often methylation-positive in high-grade disease if the hrHPV-infection had persisted longer.[16]

Ours is the first study to show significant differences in methylation, within the uniform histological diagnosis of CIN2, the outcome of which is highly variable and depends on the intrinsic progressive or regressive potential as well as clinical management philosophy. We reveal a new utility of S5 DNA-methylation measurement specifically as a classifier for assessing risk of progression in histologically confirmed CIN2.

A prognostic test for CIN could greatly alter treatment algorithms. A well-recognized dilemma of expectant management strategies is the great intra- and inter-observer variability in both cytological and histological diagnoses.[29–31] This results in misclassification of lesions, multiple follow-up visits and either delayed or premature treatment, exacerbating the potential harm to the patient. An improved predictive test could revolutionize management of CIN2 as cases with progressive potential could be treated sooner and regressive cases managed expectantly with persistent CIN2 perhaps eventually also going untreated for longer periods to allow more regressions. Additional studies on predicting the risk of progression from CIN3 to invasive cancer are also warranted. Clearly, the biological and clinical meaning of CIN2 persistence remains a major issue.

Funding

This work was supported by Sigrid Jusélius Foundation; Finnish Medical Foundation; Helsinki and Uusimaa Hospital District; Cancer Foundation (Finland); and Cancer Research UK (grant number: C569/A10404).

Conflicts of Interest

Potential conflicts of interest: JD has grants from Roche and from Genomica, outside the submitted work; KL, KA, BN, RB, MJ, IK, PN and AL no conflicts.

References

1. Castle PE, Schiffman M, Wheeler CM, Solomon D. Evidence for frequent regression of cervical intraepithelial neoplasia-grade 2. Obstet Gynecol **2009**; 113:18–25.
2. Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD. Natural history of cervicovaginal papillomavirus infection in young women. N Engl J Med **1998**; 338:423–428.
3. Massad LS, Einstein MH, Huh WK, et al. 2012 Updated Consensus Guidelines for the Management of Abnormal Cervical Cancer Screening Tests and Cancer Precursors. Obstet Gynecol **2013**; 121:829–846.
4. Tainio K, Athanasiou A, Tikkinen KAO, et al. Clinical course of untreated cervical intraepithelial neoplasia grade 2 under active surveillance: systematic review and meta-analysis. BMJ **2018**; 360:k499.
5. Kyrgiou M, Koliopoulos G, Martin-Hirsch P, Arbyn M, Prendiville W, Paraskevaidis E. Obstetric outcomes after conservative treatment for intraepithelial or early invasive cervical lesions: systematic review and meta-analysis. Lancet (London, England) **2006**; 367:489–498.
6. Arbyn M, Kyrgiou M, Simoens C, et al. Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis. BMJ **2008**; 337:a1284.
7. Kyrgiou M, Athanasiou A, Paraskevaidi M, et al. Adverse obstetric outcomes after local treatment for cervical preinvasive and early invasive disease according to cone depth: systematic review and meta-analysis. BMJ **2016**; 354:i3633.
8. Kyrgiou M, Athanasiou A, Kalliala IEJ, et al. Obstetric outcomes after conservative treatment for cervical intraepithelial lesions and early invasive disease. Cochrane database Syst Rev **2017**; 11:CD012847.

9. Moscicki AB, Ma Y, Wibbelsman C, et al. Rate of and risks for regression of cervical intraepithelial neoplasia 2 in adolescents and young women. Obstet Gynecol **2010**; 116:1373–1380.

10. Guan P, Howell-Jones R, Li N, et al. Human papillomavirus types in 115,789 HPV-positive women: A meta-analysis from cervical infection to cancer. Int J Cancer **2012**; 131:2349–2359.

11. Galgano MT, Castle PE, Atkins KA, Brix WK, Nassau SR, Stoler MH. Using biomarkers as objective standards in the diagnosis of cervical biopsies. Am J Surg Pathol **2010**; 34:1077–1087.

12. Guedes AC, Brenna SM, Coelho SA, Martinez EZ, Syrjanen KJ, Zeferino LC. p16(INK4a) Expression does not predict the outcome of cervical intraepithelial neoplasia grade 2. Int J Gynecol Cancer **2007**; 17:1099–1103.

13. Omori M, Hashi A, Nakazawa K, et al. Estimation of prognoses for cervical intraepithelial neoplasia 2 by p16INK4a immunoexpression and high-risk HPV in situ hybridization signal types. Am J Clin Pathol **2007**; 128:208–217.

14. Mirabello L, Schiffman M, Ghosh A, et al. Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia. Int J cancer **2013**; 132:1412–1422.

15. Louvanto K, Franco EL, Ramanakumar A V, et al. Methylation of viral and host genes and severity of cervical lesions associated with human papillomavirus type 16. Int J cancer **2015**; 136:E638-45.

16. De Strooper LMA, Meijer CJLM, Berkhof J, et al. Methylation analysis of the FAM19A4 gene in cervical scrapes is highly efficient in detecting cervical carcinomas and advanced CIN2/3 lesions. Cancer Prev Res (Phila) **2014**; 7:1251–7.

17. Kalantari M, Osann K, Calleja-Macias IE, et al. Methylation of human papillomavirus

16, 18, 31, and 45 L2 and L1 genes and the cellular DAPK gene: Considerations for use as biomarkers of the progression of cervical neoplasia. Virology **2014**; 448:314–321.

18. Clarke MA, Luhn P, Gage JC, et al. Discovery and validation of candidate host DNA methylation markers for detection of cervical precancer and cancer. Int J Cancer **2017**; 141:701–710.

19. Lorincz AT, Brentnall AR, Scibior-Bentkowska D, et al. Validation of a DNA methylation HPV triage classifier in a screening sample. Int J Cancer **2016**; 138:2745–2751.

20. Brentnall AR, Vasiljevic N, Scibior-Bentkowska D, et al. A DNA methylation classifier of cervical precancer based on human papillomavirus and human genes. Int J cancer **2014**; 135:1425–1432.

21. Brentnall AR, Vasiljevic N, Scibior-Bentkowska D, et al. HPV33 DNA methylation measurement improves cervical pre-cancer risk estimation of an HPV16, HPV18, HPV31 and \textit{EPB41L3} methylation classifier. Cancer Biomarkers **2015**; 15:669–675.

22. group set up by the Finnish Medical Society Duodecim, the Finnish Colposcopy Association. Cytological changes in the cervix, vagina and vulva (online). Helsinki: Finnish Medical Society Duodecim, 2016. Available at: www.kaypahoito.fi. Accessed 21 November 2016.

23. McShane LM, Altman DG, Sauerbrei W, et al. REporting recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat **2006**; 100:229–235.

24. Söderlund-Strand A, Carlson J, Dillner J. Modified general primer PCR system for sensitive detection of multiple types of oncogenic human papillomavirus. J Clin

Microbiol **2009**; 47:541–6.

25. Vasiljević N, Scibior-Bentkowska D, Brentnall AR, Cuzick J, Lorincz AT. Credentialing of DNA methylation assays for human genes as diagnostic biomarkers of cervical intraepithelial neoplasia in high-risk HPV positive women. Gynecol Oncol **2014**; 132:709–714.
26. Vasiljević N, Scibior-Bentkowska D, Brentnall A, Cuzick J, Lorincz A. A comparison of methylation levels in HPV18, HPV31 and HPV33 genomes reveals similar associations with cervical precancers. J Clin Virol **2014**; 59:161–166.
27. Brentnall AR, Vasiljević N, Scibior-Bentkowska D, et al. A DNA methylation classifier of cervical precancer based on human papillomavirus and human genes. Int J Cancer **2014**; 135:1425–1432.
28. Kelly H, Chikandiwa A, Warman R, et al. Associations of human gene EPB41L3 DNA methylation and cervical intraepithelial neoplasia in women living with HIV-1 in Africa. AIDS **2018**; :1.
29. Ismail SM, Colclough AB, Dinnen JS, et al. Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia. BMJ **1989**; 298:707–710.
30. Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. Jama **2001**; 285:1500–1505.
31. Carreon JD, Sherman ME, Guillen D, et al. CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from a histological review of population-based cervical samples. Int J Gynecol Pathol **2007**; 26:441–446.

Table 1. Baseline characteristics of the 149 women from the cohort study on expectant management of CIN2. All recruited women had a biopsy confirmed CIN2 histology at the baseline of the study. The women were divided into three clinical outcome groups (regression <CIN1; persistence CIN1/2; and progression ≥CIN3) according to the histological findings during the follow-up of the study.[a]

	All women (n=149)	**Regression** (n=88)	**Persistence** (n=36)	**Progression** (n=25)
Age (mean)	26.0	25.9	25.3	27.0
Smoking (n=128)				
No	61	41	8	11
Yes	67	37	23	7
Cig per day (mean)	5.4	5.8	4.7	5.9
Pap cytology				
NILM	23	15	7	1
≥ASC-US	126	73	29	24
Any HPV (n=141)				
Negative	25	21	2	2
Positive	116	63	32	21
HPV16+	61	32	15	14
HPV18+	11	4	5	2
HPV31+	19	8	7	4
HPV33+	8	4	3	1

[a] Significant differences (Mann-Whitney or Fisher's exact test or nonparametric test for trend) between the baseline characteristics among women in three different outcome groups was only detected with any HPV status with regression vs. persistence (p=0.02) comparison.

Abbreviations: NILM: no intraepithelial lesion or malignancy, CIN: cervical intraepithelial neoplasia, ASC-US: atypical squamous cells of undetermined significance, Cig: cigarettes; Pap: Papanicolaou

Table 2. **Odd ratios for the association between different methylation biomarkers and clinical outcome comparisons.**

Univariable odds ratios (OR, with 95% confidence intervals) for the associations between: 1) upper tertile[a] OR1, and 2) mean methylation, OR2 levels of the host gene *EPB41L3* (CpG 438, 427, 425) and viral HPV16 L1 gene (CpG 6367, 6389) and the S5 Classifier (>0.8 cut-off) and the different clinical outcome comparisons. The last column shows area under the curve (AUC) derived from receiver-operating characteristic analysis of the diagnostic performance of mean methylation cut-offs of *EPB41L3*, viral HPV16 L1 gene and the >0.8 cut-off for S5 Classifier with the different clinical outcome comparisons.

Clinical outcome comparison	**Methylation marker**	**OR1 (95% CI)[a]**	**OR2 (95% CI)**	**AUC (95% CI)**
Regression vs. persistence	EPB41L3	1.03 (0.46, 2.35)	1.06 (0.95, 1.18)	0.518 (0.41, 0.63)
	HPV16L1	1.21 (0.55, 2.65)	0.99 (0.97, 1.01)	0.497 (0.40, 0.59)
	S5 classifier	**2.61 (1.03, 6.61)**	1.04 (0.95, 1.14)	0.567 (0.46, 0.68)
Regression vs. progression	EPB41L3	2.29 (0.78, 6.68)	**1.14 (1.03, 1.26)**	**0.649 (0.52, 0.77)**
	HPV16L1	1.92 (0.79, 4.73)	1.01 (0.99, 1.03)	0.576 (0.46, 0.69)
	S5 classifier	**4.84 (1.35, 17.41)**	**1.17 (1.06, 1.30)**	**0.718 (0.61, 0.83)**
Persistence vs. progression	EPB41L3	2.55 (0.78, 8.34)	1.08 (0.99, 1.19)	**0.639 (0.50, 0.78)**
	HPV16L1	1.59 (0.57, 1.44)	**1.03 (1.00, 1.07)**	0.588 (0.45, 0.73)
	S5 classifier	2.86 (0.88, 9.33)	**1.15 (1.01, 1.30)**	**0.676 (0.54, 0.82)**
Regression/persistence vs. progression	EPB41L3	2.28 (0.80, 6.49)	**1.12 (1.03, 1.21)**	**0.646 (0.52, 0.77)**
	HPV16L1	1.82 (0.77, 4.33)	1.01 (0.99, 1.03)	0.580 (0.46, 0.70)
	S5 classifier	**4.48 (1.27, 15.77)**	**1.16 (1.06, 1.28)**	**0.706 (0.60, 0.81)**

Regression vs. persistence/progression	EPB41L3	1.37 (0.68, 2.75)	**1.09 (1.01, 1.19)**	0.572 (0.48, 0.67)
	HPV16L1	1.47 (0.76, 2.83)	1.00 (0.98, 1.02)	0.530 (0.44, 0.62)
	S5 classifier	**2.68 (1.27, 5.64)**	**1.10 (1.02, 1.19)**	**0.630 (0.54, 0.72)**

[a] In OR1 a high tertile level was defined as 1/3 of the upper methylation levels that was identified in each of the outcome category comparisons.

AUC: Area under the ROC curve. Significant ORs are shown in bold.

Table 3. Odds ratios for the association between the different clinical outcome comparisons and the different markers.

Comparison between: 1) upper tertile[a] level of S5 Classifier, 2) Pap cytology comparison of ≤ASC-US vs ≥LSIL, 3) HPV 16/18 and 4) HPV 16/18/31/33 genotyping positive or negative. S5 and HPV16/18/31/33 genotyping were the two significant prognostic variables, shown in bold.

Clinical outcome	**OR (95% CI)**			
	S5 Classifier	**Pap cytology ≤ASC-US vs ≥LSIL**	**HPV16/18 genotyping**	**HPV16/18/31/33 genotyping**
Regression	1.00	1.00	1.00	1.00
Persistence	1.33 (0.58, 3.07)	1.00 (0.43, 2.33)	1.99 (0.91, 4.35)	**3.50 (1.44, 8.52)**
Progression	**3.39 (1.35, 8.50)**	2.32 (0.73, 7.42)	2.38 (0.96, 5.91)	**3.17 (1.15, 8.68)**

[a] A high tertile level was defined as 1/3 of the upper methylation levels of S5 Classifier at baseline.

Figure Legends

Figure 1. Flow chart of the expectant management of the cervical intraepithelial neoplasia (CIN) grade 2-study.

The study (ISRCTN91953024) started in the colposcopy clinic of Helsinki University Hospital, Finland in September 2013 and is ongoing. The flow chart shows the numbers for the first 149 young (18-30-year-old) women included in the current analyses who had a minimum of two follow-up visits completed as of November 2017. Women with follow-up diagnoses less than CIN grade 1 (<CIN1) were categorized as regression, women with CIN1 and/or CIN2 were categorized as persistence and women with CIN grade 3 or worse (≥CIN3) were categorized as progression. Women with histological ≥CIN3 were treated by loop electrosurgical excision procedure (LEEP), as were all women with a diagnosis (histological or colposcopic) of CIN1 or CIN2 after two years.

Figure 2. Receiver-operating characteristic (ROC)-curves for the performance of S5 classifier and in combination with other tests.

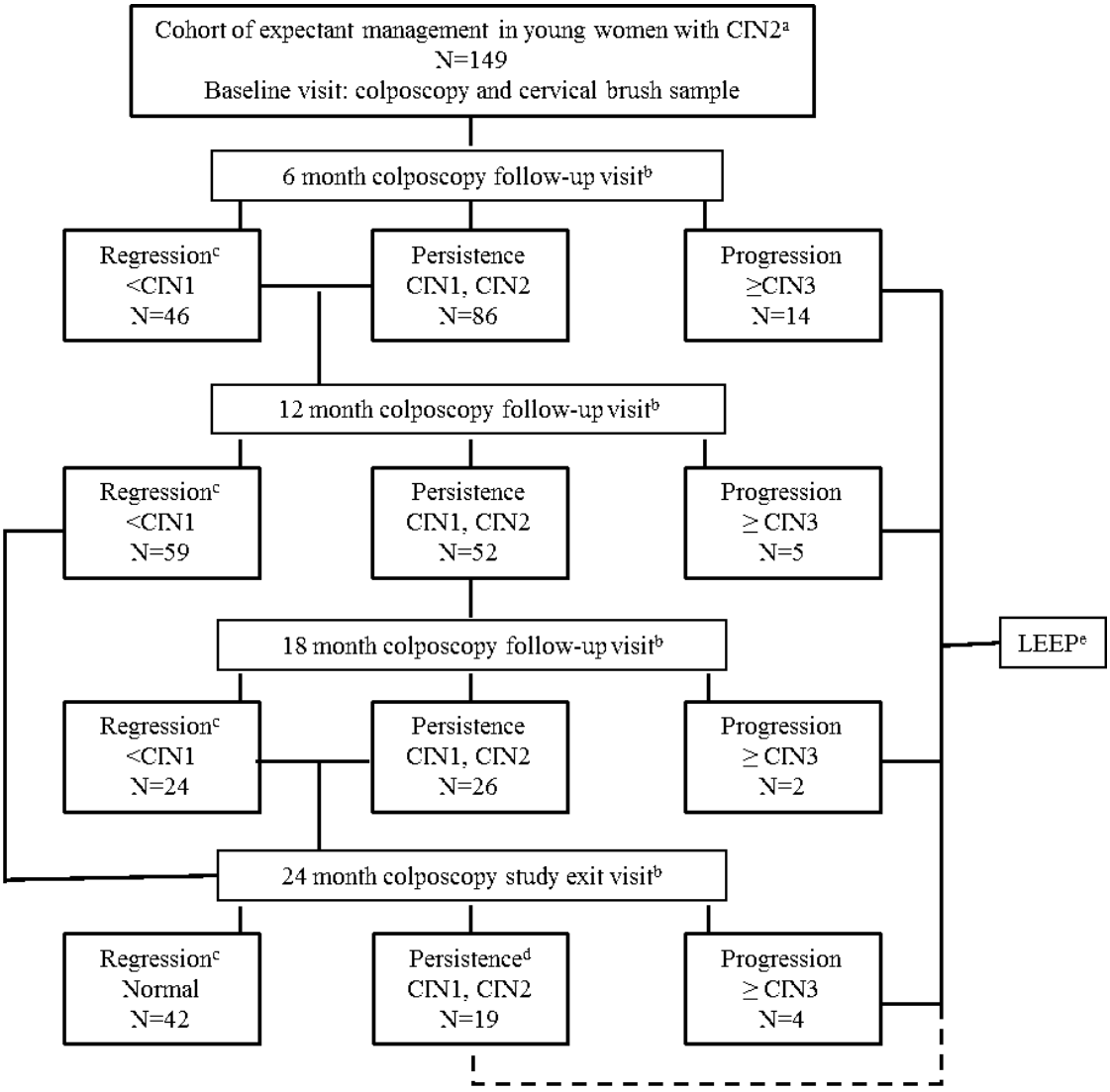
S5 classifier performance alone and in combination with ≥HSIL cytology and/or HPV16/18 or HPV16/18/31/33 genotyping positivity tested in different clinical outcome categories. S5 classifier alone (red dashed line, AUC 1), S5 combined with ≥HSIL (black solid line, AUC 2), S5 combined with ≥HSIL and HPV16/18 (green dotted line, AUC 3) or S5 combined with ≥HSIL and HPV16/18/31/33 (orange dash line, AUC 4) in the following clinical outcome categories: A) regression vs. progression; B) regression/persistence vs. progression; C) regression vs. persistence/progression. For each clinical outcome category (A-C) the corresponding performance of the single tests: HPV16/18 genotyping (black solid circle), HPV16/18/31/33 genotyping (hollow black circle) and the different cytological abnormality end points are shown; ≥ASC-US (hollow diamond), ≥LSIL (hollow triangle) and ≥HSIL (hollow square).[a]

Abbreviations: AUC: Area Under the ROC Curve, ASC-US: atypical squamous cells of undetermined significance, LSIL: Low-grade squamous intraepithelial lesion, HSIL: High-grade squamous intraepithelial lesion

Figure 3. Cumulative proportions of women who progressed to ≥CIN3 by time since the diagnosis of CIN2.

In this analysis persistent CIN1 or CIN2 were regarded as nonprogressions. The graph shows the distribution by time (in months) of women positive for the following: S5, HPV16/18, and cytology ≥HSIL (solid line) versus women who were negative for all these markers: S5 Classifier ≤0.8, HPV16/18 negative and pap-smear <HSIL (dashed line). There was a significant difference between these predictors (LR-test *p=0.03*).

Figure 1



[a] Inclusion criteria: Histological diagnosis of CIN2, 18-30 years of age, non-pregnant, adequate colposcopy (type I or II transformation zone, lesion size <¾ of the transformation zone, written informed consent; Exclusion criteria: Prior treatments for cervical intraepithelial neoplasia (CIN), vulvar intraepithelial neoplasia (VIN), and vaginal intraepithelial neoplasia (VAIN) or corresponding cancers, known HIV positivity, immunosuppressive medication, lactation, no common language

[b] Colposcopy, cytology, punch biopsies, and cervical brush sample

[c] In cases of histological regression, but high grade cytology (atypical squamous cells cannot rule out HSIL (ASC-H) or high grade squamous intraepithelial lesion (HSIL)) lesions were considered persistent

[d] In cases of persisting CIN1 decisions to treat were based on individual assessment taking into account also the cytological and colposcopic findings

[e] Loop electrosurgical excision procedure under local anesthesia in out-patient colposcopy clinic

Figure 2

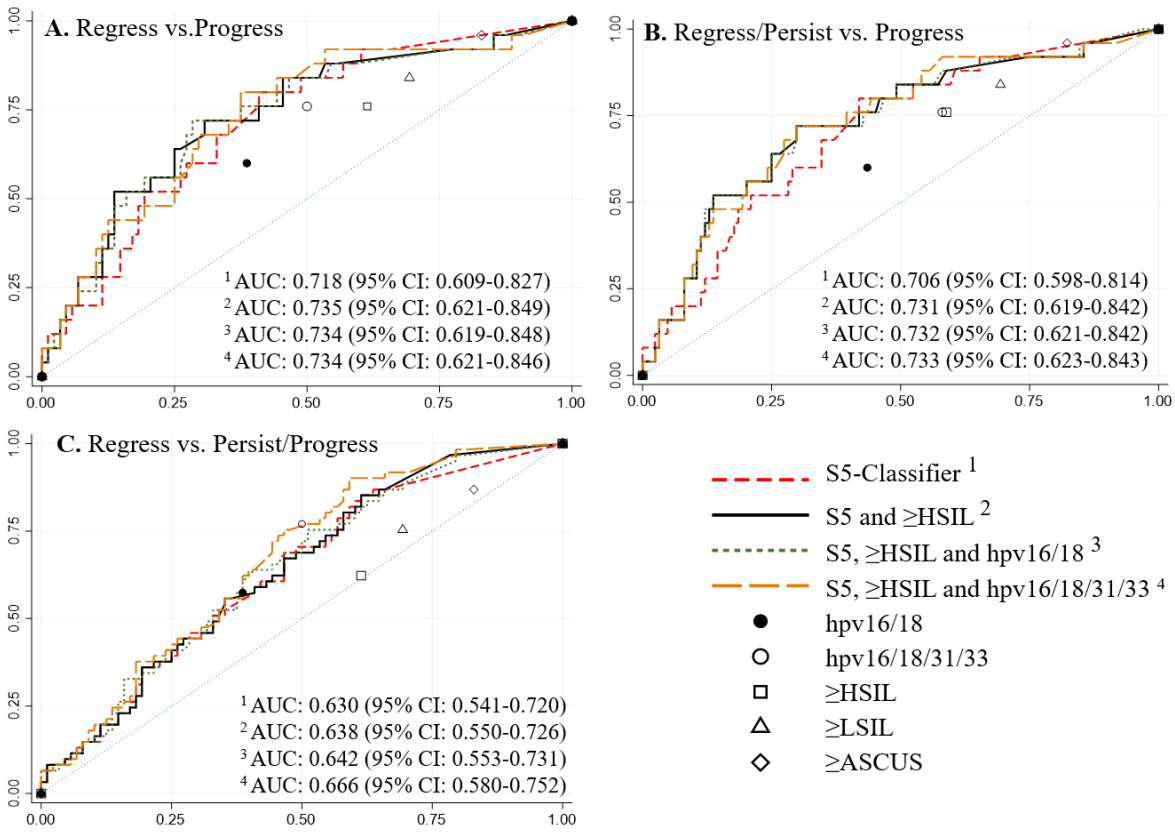


Figure 3