IJC
International Journal of Cancer

Performance of a new HPV and biomarker assay in the management of hrHPV positive women: Subanalysis of the ongoing multicenter TRACE clinical trial ($n > 6{,}000$) to evaluate POU4F3 methylation as a potential biomarker of cervical precancer and cancer

Adrienn Kocsis[1,2], Tibor Takács[1,2,3], Csaba Jeney[4], Zsuzsa Schaff[5], Róbert Koiss[6], Balázs Járay[5], Gábor Sobel[7], Károly Pap[8], István Székely[6], Tamás Ferenci[9], Hung-Cheng Lai[10,11], Miklós Nyíri[1,2] and Márta Benczik[1,2,3]

[1] NEUMANN Diagnostics Ltd., Budapest, Hungary
[2] Cellcall Ltd., Budapest, Hungary
[3] SYNLAB Hungary Ltd., GenoID Molecular Diagnostic Laboratory, Budapest, Hungary
[4] Institute of Medical Microbiology, Semmelweis University, Budapest, Hungary
[5] Second Department of Pathology, Semmelweis University, Budapest, Hungary
[6] Department of Obstetrics and Gynaecology, St. Stephan Hospital, Budapest, Hungary
[7] Second Department of Obstetrics and Gynaecology, Semmelweis University, Budapest, Hungary
[8] Jósa András Regional Central Hospital, Nyíregyháza, Hungary
[9] Óbuda University, Budapest, Hungary
[10] Department of Obstetrics and Gynecology, Shuang Ho Hospital, Taipei Medical University, Taipei, Taiwan
[11] Department of Obstetrics and Gynecology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

The ongoing Triage and Risk Assessment of Cervical Precancer by Epigenetic Biomarker (TRACE) prospective, multicenter study aimed to provide a clinical evaluation of the CONFIDENCE™ assay, which comprises a human papillomavirus (HPV) DNA and a human epigenetic biomarker test. Between 2013 and 2015 over 6,000 women aged 18 or older were recruited in Hungary. Liquid-based cytology (LBC), high-risk HPV (hrHPV) DNA detection and single target host gene methylation test of the promoter sequence of the POU4F3 gene by quantitative methylation-specific polymerase chain reaction (PCR) were performed from the same liquid-based cytology sample. The current analysis is focused on the baseline cross-sectional clinical results of 5,384 LBC samples collected from subjects aged 25 years or older. The performance of the CONFIDENCE HPV™ test was found to be comparable to the cobas® HPV test with good agreement. When applying the CONFIDENCE Marker™ test alone in hrHPV positives, it showed significantly higher sensitivity with matching specificity compared to LBC-based triage. For CIN3+ histological endpoint in the age group of 25–65 and 30–65, the methylation test of POU4F3 achieved relative sensitivities of 1.74 (95% CI: 1.25–2.33) and 1.64 (95% CI: 1.08–2.27), respectively, after verification bias adjustment. On the basis of our findings, POU4F3 methylation as a triage test of hrHPV positives appears to be a noteworthy method. We can reasonably assume that

Key words: cervical cancer, high-risk HPV, POU4F3 biomarker, epigenetics, host gene methylation

Abbreviations: AGC: atypical glandular cells; ASCH: atypical squamous cells-cannot exclude HSIL; ASCUS: atypical cells of undetermined significance; CC: cervical carcinoma; CI: confidence intervals; CIN: cervical intraepithelial neoplasia; CIN2+: cervical intraepithelial neoplasia grade two or worse; CpG: C-phosphate-G; CV: coefficient of variation; HPV: human papillomavirus; hrHPV: high-risk HPV; HSIL: high-grade squamous intraepithelial lesion; IQR: interquartile range; LBC: liquid-based cytology; LSIL: low-grade squamous intraepithelial lesion; NILM: negative for intraepithelial lesion or malignancy; NPA: negative percent agreement; PABAK: prevalence-adjusted and bias-adjusted kappa value; PCR: polymerase chain reaction; PPA: positive percent agreement; qMSP: quantitative methylation-specific real-time PCR; ROC: receiver operating characteristic; SD: standard deviation; TRACE: triage and risk assessment of cervical precancer by epigenetic biomarker; VIA: acetic acid test

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: New Hungary Development Plan, Hungarian Government; **Grant number:** KMR_12-1-2012-0032; **Grant sponsor:** SYSTEMCERV, FP7 SME-targeted Collaborative Project, EU; **Grant number:** HEALTH-2012.2.1.2-1; **Sponsor:** Cellcall Ltd., Budapest, Hungary; **Sponsor:** NEUMANN Diagnostics Ltd., Budapest, Hungary
DOI: 10.1002/ijc.30534
History: Received 27 Apr 2016; Accepted 8 Nov 2016; Online 22 Nov 2016
Correspondence to: NEUMANN Diagnostics Ltd., Röppentyű u. 48, Budapest H-1139, Hungary, Tel.: +36-70-454-3251, E-mail: adrienn.kocsis@neumanndx.com

its quantitative nature offers the potential for a more objective and discriminative risk assessment tool in the prevention and diagnostics of high-grade cervical intraepithelial neoplasia (CIN) lesions and cervical cancer.

What's new?

Combining DNA methylation biomarker detection with high-risk human papillomavirus (hrHPV) testing is a promising screening option for cervical cancer. Currently, one of the largest clinical studies designed to evaluate human epigenetic biomarker testing for cervical screening is the Triage and Risk Assessment of Cervical Precancer by Epigenetic Biomarker (TRACE) study. In this TRACE analysis, methylation of the *POU4F3* promoter, a candidate marker for high-grade HPV-positive cervical intraepithelial lesions (CIN3+), showed significantly higher sensitivity and similar specificity for CIN3+ than liquid-based cytology. The findings suggest that quantitative methylation of *POU4F3* is a valuable tool for high-grade CIN detection.

Cervical cancer is by far the most common human papillomavirus (HPV)-related disease, resulting in 270,000 women deaths annually worldwide.[1]

Cytology-based screening was successful in the developed countries due to organized screening programs. However, cytology has several limitations like subjectivity, interobserver variability, modest sensitivity, indication of unnecessary treatment and cost burdens on the healthcare system.[2–4]

The newest guidelines recommend primary high-risk HPV (hrHPV) screening as an alternative to cytology-based cervical cancer screening.[5–7] Molecular hrHPV detection has several benefits compared to cytology being more reproducible and reliable, which enables objective evaluation with the possibility of quality-controlled and automated high-throughput application.[3,8,9] Although HPV testing as a viral marker offers the best negative predictive value for high-grade cervical intraepithelial neoplasia (CIN) lesions with long-term confidence, further aspects are to be considered when used in the clinical setting.[5,10–14] Although hrHPV detection has high sensitivity, because of its moderate specificity, HPV testing cannot be applied for screening alone, but triage methods are required. Additionally, the adequate risk stratification of hrHPV positive women is not yet clear.[4,15,16]

The most recent guidelines recommend different triage strategies including cytology alone as a reflex test[6] or combination of genotyping for HPV16/18 and cytology[5] after primary hrHPV screening. At the same time, new molecular biomarkers were proposed for the management of hrHPV positive women. These alternatives are currently under evaluation including immunostaining, host methylation, viral methylation testing and detection of microRNA expression.[15] The DNA methylation of C-phosphate-G (CpG) islands in promoter regions of tumor suppressor genes was also shown to be a potential biomarker for early cancer detection. DNA methylation of various host cell genes has been detected in cervical cancer and precancer.[17–19] In concurrent or sequential screening strategies the combination of DNA methylation and a highly sensitive test such as HPV detection, may be a reasonable screening option.[20] The methylation status of different genes, including CADM1, MAL and PAX1,[20–27] was tested as possible triage options in hrHPV positive women.

The methylation of gene POU4F3 (POU Class 4 Homeobox 3) was described as a potential molecular triage tool in HPV positives for cervical intraepithelial lesion grade three or worse (CIN3+) by Pun *et al.*[28] POU4F3 showed 74% sensitivity and 89% specificity, which represented the best performance compared to other evaluated biomarkers.[28] However, the combined use of hrHPV assay and POU4F3 methylation marker warrants further investigations.[28]

The Triage and Risk Assessment of Cervical Precancer by Epigenetic Biomarker (TRACE) ongoing clinical trial evaluated the host gene methylation of POU4F3 as a cervical biomarker and as a potential triage method of hrHPV positive women. A complex molecular assay was applied in the study, which comprises a newly developed HPV DNA test and a human epigenetic biomarker test.

Materials and methods

Objectives, study design and interventions

The TRACE prospective, multicenter clinical study aimed to assess a quantitative methylation-specific PCR (qMSP) assay designed to detect the promoter sequence methylation of the gene POU4F3 as the candidate test compared to cytology-based triage as the reference test in HPV positive women. The aim of our study was to validate the CONFIDENCE™ assay on >6,000 cervical samples evaluating the clinical performance of the CONFIDENCE Marker™ as a potential host methylation-based triage test of hrHPV positives.

The study was intended for noninferiority in specificity assuming 90% specificity of the reference test with a lower confidence bound not exceeding 85%, whereas a superiortiy design was used to compare sensitivity assuming 60% sensitivity[3] of the reference test and supposing that the candidate test is able to reach 80% sensitivity. To detect the established performance with a significance level of 5% and power of 90%,[29] the required sample size was calculated to be 194 for cervical intraepithelial neoplasia grade two or worse (CIN2+) cases and 1,565 for patients without CIN2+. Assuming 16.3% CIN2 positivity rate among hrHPV positive patients,[30] the total required sample size of hrHPV positives was estimated to be 1,869.

Between 2013 and 2015 the study recruited 6,761 cervical samples collected from subjects over 18 years of age

including 1,685 hrHPV positives. The follow-up phase of the study is ongoing. The liquid-based cytology (LBC) samples were collected from women aged between 18 and 65 years, who underwent cervical sampling at one of the five participating clinical sites.

Four sites participating in the study were outpatient clinics. The characteristics of samples received from these four sites were similar (Supporting Information Table S3). The women visiting these sites constituted the outpatient population of the study; they were screened and followed-up according to the Hungarian guidelines.[31] In their case, LBC sampling was performed followed by colposcopy-assisted visual inspection of the cervix with acetic acid test (VIA). Depending on the result of VIA, LBC and HPV detection gynecologists made medical decisions and if necessary referred these women for cone or punch biopsy. In the Hungarian screening guidelines there is no referral to colposcopy *per se*. Accordingly, this population is lacking also from our study, but at the same time all women underwent colposcopy-assisted VIA.

The rest of the patients constituted the oncology center population. These women visited a site which as a regional oncology center manages patients referred for treatment. At this site, LBC sampling was taken prior to the intervention (cone biopsy or hysterectomy). The sample characteristics of this population were different compared to the outpatient population (Supporting Information Table S3). Not all patients of the oncology center were referred by one of the four outpatient clinics and likewise, several patients were not referred and were treated by the outpatients clinics.

Since each test was performed on each sample enrolled, we were able to estimate how the different strategies would have performed in a triage setting by *post hoc* analysis. Confirmation was based on the diagnostic test results in line with the clinical protocols.

In order to assess the clinical performance of the CONFIDENCE™ assay, histology was considered as the gold standard confirmation method. In selected cases, negative confirmation was accepted based on the result of a subsequent LBC cytology and HPV test performed between 5 and 24 months (mean = 9.9 months) after the baseline sampling (*i.e.*, negative HPV test result with regression or no progression in previously negative for intraepithelial lesion or malignancy [NILM], atypical cells of undetermined significance [ASCUS] or low-grade squamous intraepithelial lesion [LSIL] cases). The study design (Fig. 1) accomodated the precondition of the Ethics Committee approval, which required that the patients were managed as per the applicable clinical guidelines and protocols.[31] Due to this, blinded referral for histology was not feasible.

The eligibility criteria for the baseline data analysis excluded subjects who did not fulfill the age limit, cases with invalid or absent result of any of the applied tests, if LBC sample was collected >12 months before the respective cervical cone or punch biopsy, or if the patient was treated within 12 months prior to the baseline sampling.

Cervical sample collection and evaluation

The LBC samples were collected by Cervex Brush®Combi (Rovers®, Oss, Netherlands) in ThinPrep® PreservCyt®Solution (Hologic®, Marlborough, MA). LBC (ThinPrep®, Hologic®, Marlborough, MA), hrHPV detection (CONFIDENCE HPV™, NEUMANN Diagnostics Ltd, Budapest, Hungary; cobas® HPV, Roche, Branchburg, NJ; Full Spectrum HPV, Synlab GenoID Laboratory, Budapest, Hungary) and biomarker test (CONFIDENCE Marker™, NEUMANN Diagnostics Ltd., Budapest, Hungary) were performed using the same LBC sample in all cases irrespective of the HPV status.

VIA was performed according to the international standards.[32] LBC was performed and evaluated respecting the quality assurance and quality control protocols and in line with the current international standards blinded to the HPV detection result and using the Bethesda system 2001[33] for reporting the result.

Technical reliability was ensured by using liquid-based cytology sampling technique with automated slide processing using ThinPrep®2000 Processor (Hologic®, Marlborough, MA) unit. All involved gynecologists were trained for LBC sampling, while all cytotechnologists and pathologists were skilled in processing the LBC samples and evaluating the slides. LBC slide preparation and clinical evaluation were conducted at two clinical laboratories. Thirty percent of the negative slides, which were selected randomly, were rescreened via rapid manual review performed by a qualified supervisory cytotechnologist.

The cone or punch biopsy samples were evaluated by the local pathologist of each clinical sites, no consensus review was applied.

CONFIDENCE™ assay

The CONFIDENCE™ assay developed by NEUMANN Diagnostics offers quality controlled high-throughput and highly automated protocols. The assay comprises the CONFIDENCE HPV™ and the CONFIDENCE Marker™ tests (NEUMANN Diagnostics Ltd., Budapest, Hungary). The CONFIDENCE HPV™ is a viral DNA test, based on multiplex real-time PCR. The CONFIDENCE Marker™ is a human epigenetic biomarker test, which measures the methylation level of a single target gene by quantitative methylation-specific real-time PCR (qMSP).

The CONFIDENCE HPV™ is a TaqMan®-based L1 region-specific multiplex real-time PCR assay for viral DNA detection. The test detects HPV16 and HPV18 separately and other high-risk types (HPV31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68) in group. The DNA extraction for HPV detection was performed on 200 µl LBC sample applying a silica-filter-plate-based purification method on the Tecan EVO® liquid handling platform using standardized, contamination safe workflow with high reliability. After the 96-channel

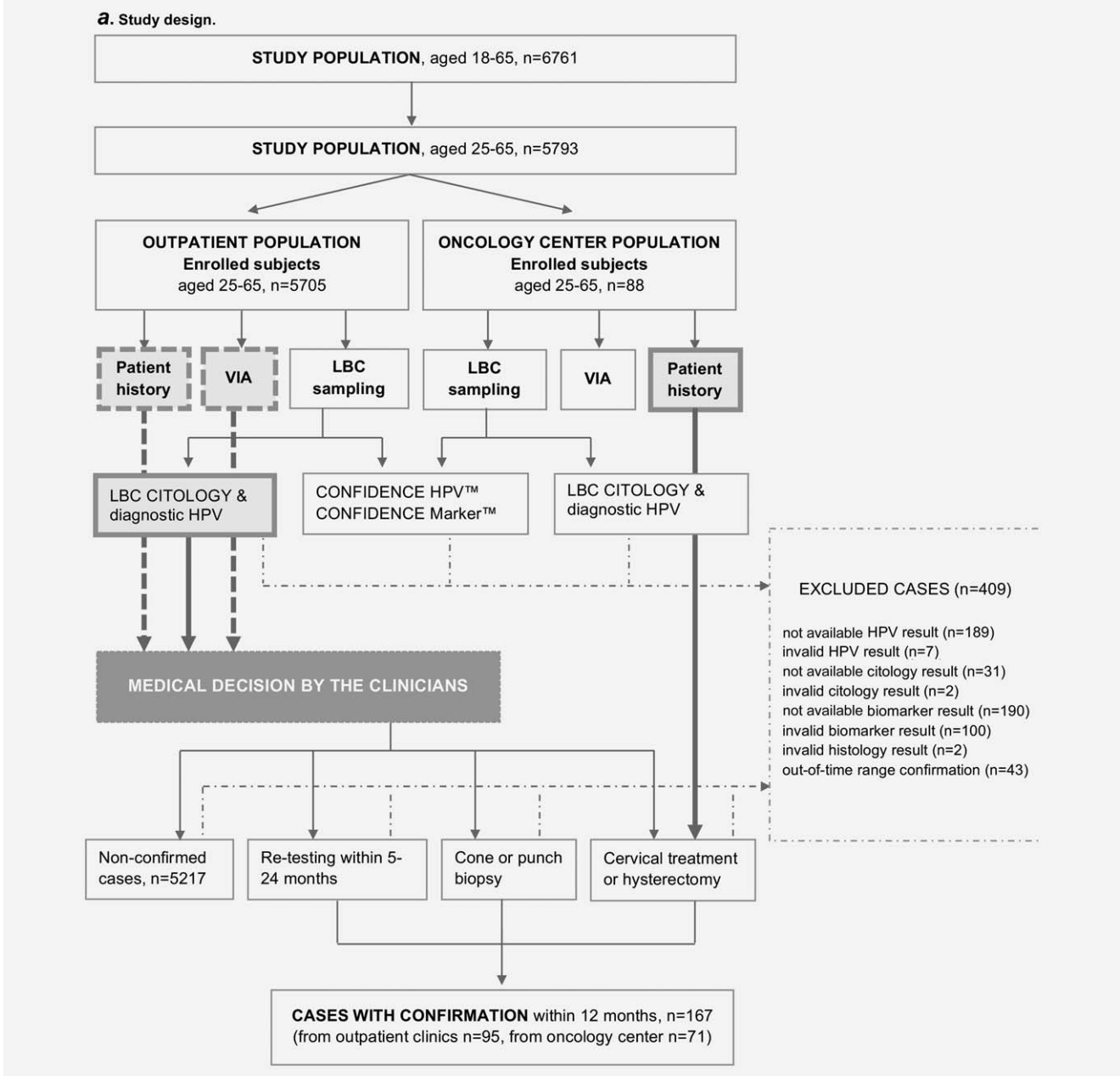

Figure 1. TRACE study patient distribution diagram (*a*, *b*). The TRACE study consists of an outpatient and an oncology center population. The medical management of the women enrolled was based on patient history, VIA, cytology and the result of diagnostic HPV tests as per the Hungarian clinical guidelines. The distribution of the CONFIDENCE™ assay test results are shown for the per protocol study population.

automated PCR setup, 5-plex quantitative real-time PCR was performed on the QuantStudio™ 6 Flex platform in 384-well plate format in four reactions per sample. The input DNA volume was 10 µl per reaction well. The sample quality was assured by the amplification of cellularity control in each sample. The DNA preparation process was controlled by the amplification of an artificial internal DNA control sequence added to the sample during the process.

The CONFIDENCE Marker™ test measures the methylation level of CpG sites in the promoter region of POU4F3 by qMSP. The quantitative measurement of the gene COL2A1 (type II collagen) is used as internal reference to normalize the methylation level of the POU4F3. The applied M-index (methylation index) calculation was proposed by Huang *et al.*[34] The M-index value ranges from 1 to 10,000 where the unit of M-index refers to a methylated ratio of 0.01% of the cells. The direct bisulfite treatment was performed on the LBC sample in 96-well plate format by using the EZ-96 DNA Methylation-Direct™ Kit (Zymo Research Co., Irvine, CA, ref. no.: D5023) according to the instructions of the manufacturer. The PCR

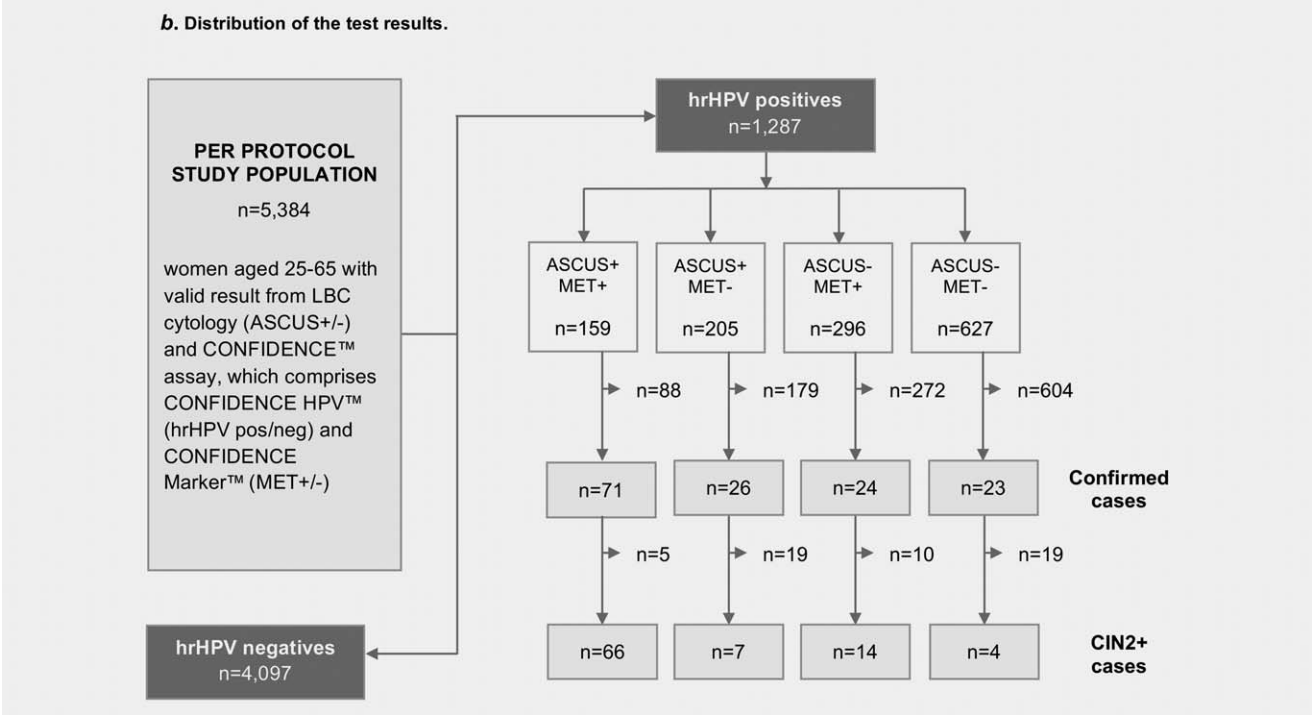

Figure 1. TRACE study patient distribution diagram (*a*, *b*). (Continued)

setup was implemented using the Tecan EVO® (Tecan, Männedorf, Switzerland) liquid handling system with standardized, highly reliable workflow. The amplification of the recovered bisulfite converted genomic target DNA and the internal reference sequence were performed using a TaqMan® probe system on the QuantStudio™ 6 Flex Real-Time PCR platform (LifeTechnologies Co., Carlsbad, CA) in 384-well plate format. The input DNA volume was 5 µl per reaction well. The sample quality was controlled by the amplification of the COL2A1 internal reference. The applied primer and probe system were previously described by Chen *et al.*[35]

The optimal cut-off value of POU4F3 biomarker positivity was established within the enrolled sample population aiming to achieve a predefined sensitivity level. The predefined sensitivity threshold level was determined based on the result of the PALMS study.[36] The Ct values of the qMSP reactions were determined and evaluated using the QuantStudio™ Real-Time PCR Software v1.2 (LifeTechnologies Co., Carlsbad, CA). If the Ct value for COL2A1 exceeded 31 and the value of M-index was below the cut-off value, the result of the methylation test was classified as detection failure.

The reproducibility of the methylation marker test was evaluated assessing Ct measurements of control DNA and randomly selected clinical samples, where the total coefficient of variation (CV) was between 1.1 and 3.8% (Supporting Information Fig. S5).

Diagnostic HPV tests

In the TRACE study the diagnostic HPV test results were provided by two tests: the first 60% of the samples were tested by cobas® HPV test (Roche, Branchburg, NJ),[37] while the remaining 40% was tested by the Full Spectrum HPV test[38] (SYNLAB GenoID Laboratory, Budapest, Hungary) in order of receipt.

The real-time PCR-based cobas® HPV test detects the HPV DNA of HPV16, 18 and other hrHPV types in a pooled manner. It was applied in line with the user manual of the system.

The Full Spectrum HPV test[38] (SYNLAB GenoID Laboratory, Budapest, Hungary) detects the high-risk and low-risk HPV types in group and subsequently performs genotyping via type-specific probes. Other HPV types are detected as pooled result without genotyping.

For study purposes, the cobas® HPV test was selected as a clinically validated comparator test to evaluate the CONFIDENCE HPV™ test, while the Full Spectrum HPV test provided diagnostic results only.

Statistics and data analysis

To assess the clinical performance, four triage strategies were compared including POU4F3 methylation alone, LBC cytology alone and both in combination with HPV genotyping. The comparison of the clinical performance of different triage methods was evaluated in the outpatient population and the oncology center population separately (Supporting Information Table S6). Since the rate of the negative cases was very low in the latter, only crude sensitivity could be meaningfully calculated in the oncology center population (Supporting Information Table S6). In the main analysis of the

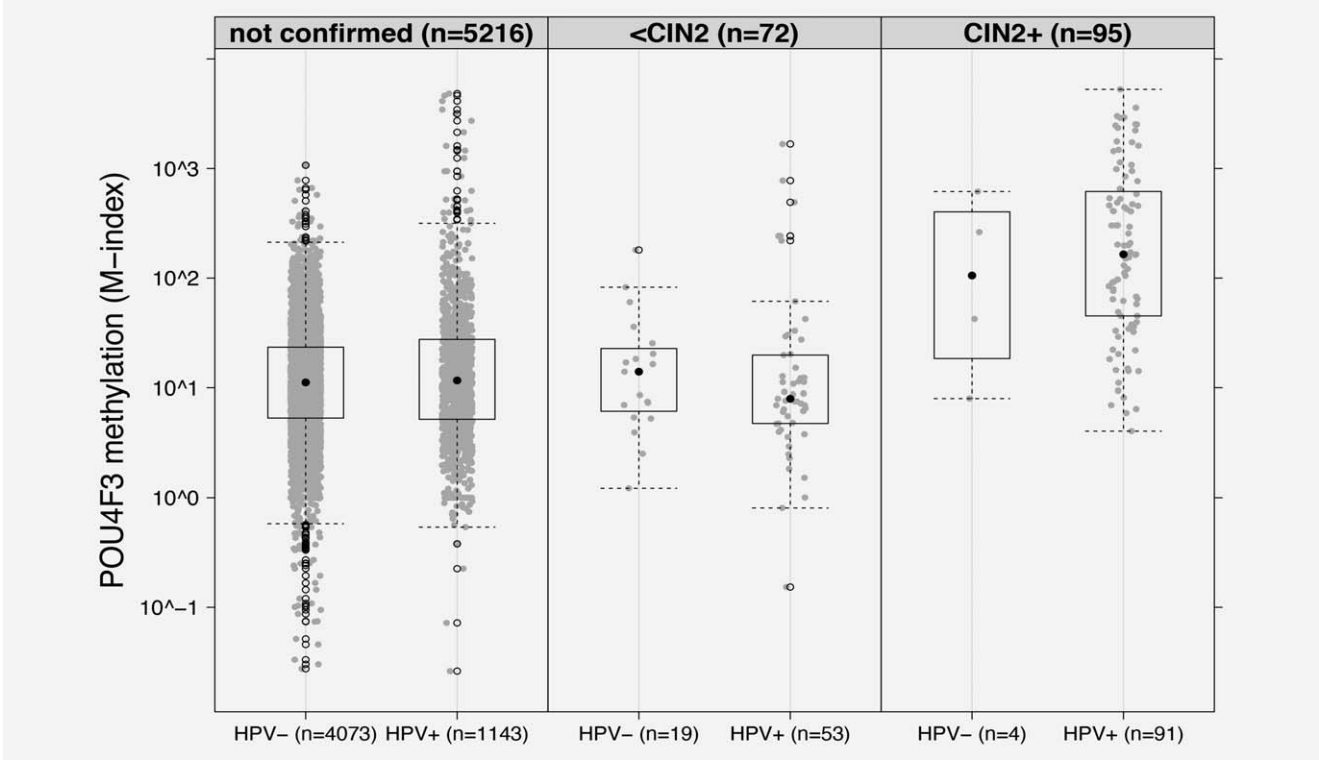

Figure 2. M-index levels by hrHPV positivity, women aged 25–65. Distribution of POU4F3 methylation (Mindex) level in different subpopulations: cases with no confirmation, confirmed <CIN2 and CIN2+.

study, the total population of the enrolled women was evaluated collectively (Table 2).

Absolute and relative sensitivity and specificity were calculated for the study endpoints of histologically confirmed high-grade CIN lesions, *i.e.*, CIN2 or worse (CIN2+) and CIN3 or worse (CIN3+), respectively. The statistical analysis was performed mainly in the age group of 25–65 and partially in the age groups of 25–29 and 30–65 years.

Categorical variables are presented as count (%), continuous variables are presented as mean (±standard deviation [SD]) and/or median (interquartile range [IQR]) [min]-[max]. The comparison of different HPV tests was evaluated by overall agreement, positive (PPA, Chamberlain's PPA) and negative percent agreement (NPA, Chamberlain's NPA),[39,40] prevalence-adjusted and bias-adjusted kappa value (PABAK) and Cohen's kappa.[41] The McNemar's test was used to test for the presence of bias.[41]

Receiver operating characteristic (ROC) and area under the ROC curve (AUC) analysis were performed to determine the sensitivity and specificity of the POU4F3 methylation in the whole sample population and in the hrHPV positive subpopulation as well.

As mentioned before, the M-index cut-off was set to achieve a sensitivity matching a predefined value, which also determined the level of specificity. Since the same sample set was used for both tuning and testing the cut-off threshold of the candidate marker, in order to establish the unbiased value of relative and absolute sensitivity and specificity, bootstrap and cross-validation were applied in data analysis. More specifically, nonparametric bootstrap resampling was used to avoid overfitting with 1,000 replicates to provide point estimates with 95% confidence intervals (CI).[42] In addition, five-fold cross-validation was used with 100-times repetition to confirm point estimates with high confidence[43] (Supporting Information Table S5).

As subjects were mainly referred for verification on the basis of the positive test result, verification rate was highly different between test-positive and test-negative cases. Ignoring this aspect, *i.e.*, assessing merely crude estimates based on subjects with case confirmation, would have led to substantial verification bias. Therefore, verification bias adjustment was performed using the weighted generalized estimating equations approach described by Xue *et al.*[44] In both the crude and verification bias-adjusted estimations, absolute and relative parameters were calculated with 95% CIs.

Statistical analysis was carried out using R statistical software (v0.99.491; http://www.R-project.org), FileMaker®Pro database management software (13.0v5, FileMaker) and MATLAB® R2010a (v7.10; The MathWorks).

Results

Study characteristics

A total of 6,215 women aged 18 or older with valid test results were enrolled into the study (Fig. 1). The mean age in the whole population was 35.8 (±9.9) years, and most of the subjects were aged between 25 and 40 ($n = 3{,}505$). Out of the

5,793 baseline samples collected from women aged 25–65 overall 5,384 had valid test results, which were included in the current cross-sectional baseline subanalysis (Table 1a). The test results of the baseline samples belonging to women aged 18–65 are also presented (Supporting Information Table S4).

Overall 23.9% ($n = 1,287$) of the LBC samples was hrHPV positive by the CONFIDENCE HPV^{TM} test. The prevalence was 7% ($n = 376$) for HPV16, 2% ($n = 109$) for HPV18, and 18.4% for other hrHPV types ($n = 993$), respectively, while 14.5% ($n = 186$) of the infections were multiplex (Table 1a). Overall 86% ($n = 4,628$) of the LBC specimens were classified as negative for NILM. The prevalence of abnormal cytology is detailed in Table 1a. The rate of borderline cases (*i.e.*, ASCUS and LSIL) was 12.5% ($n = 665$), while 1.7% ($n = 91$) of the subjects had high-grade lesions (*i.e.*, high-grade squamous intraepithelial lesion [HSIL], atypical squamous cells-cannot exclude HSIL [ASCH], cervical carcinoma [CC] and atypical glandular cells [AGC]).

As it is demonstrated in Table 1a, the median M-index level showed age-related distribution: 8 (12.1), 9.8 (16.3), 14 (22.3), 20.6 (33.4) and 21.5 (28,8) for the corresponding age groups, respectively.

Overall 139 of the enrolled subjects were confirmed histologically following the clinical protocol based on the international standards,[45] *i.e.*, 28 negative, 111 positive including 95 CIN2+, 77 CIN3+ (Table 1a).

In addition, 28 samples with no histology were accepted as CIN2 negatives based on follow-up. Therefore, the total number of samples with valid disease verification was 167 (*i.e.*, 56 negative, 111 positive). Out of the total 167 confirmed cases 144 (86.2%) were hrHPV positive, 108 (64.7%) were ASCUS+ and 95 (56.9%) were evaluated as biomarker positive (MET+). The methylation index of POU4F3 was elevated in all of the 12 histologically confirmed invasive or *in situ* carcinoma cases (Table 1b).

Validation of the CONFIDENCE HPV^{TM} test

In the samples colllected from women over 25 years of age, the CONFIDENCE HPV^{TM} was compared to cobas® HPV on 3,150 samples resulting in 92.3% overall agreement. PPA was 83.3% (95% CI: 81.2–85.4) and NPA was 95% (95% CI: 94.4–95.6). Agreement levels for HPV16, HPV18 and other hrHPV types were comparable. The value of PABAK was 0.85 (95% CI: 0.83–0.86), while Cohen's kappa was 0.78 (95% CI: 0.75–0.82). The clinical sensitivity and specificity for CIN2+ cases were found to be equivalent for the CONFIDENCE HPV^{TM} and the cobas® HPV tests (Table 1c).

Characteristics and optimization of POU4F3 biomarker positivity

The distribution of M-index value by hrHPV status and confirmatory results (*i.e.*, no confirmation, <CIN2, CIN2+) are presented in Figure 2. In the age group over 25 years the median value of M-index was 11.1 (18.0) in the hrHPV negative NILM cases, where the median age was 37.3 (12.3). Among the hrHPV positive NILMs the median value of M-index was 11.9 (22.5) with median age of 33.2 (10.7) (Fig. 2). Since the median value of methylation level was similar, we can reasonably assume that the methylation of POU4F3 is independent from hrHPV positivity, which can be a major advantage in the triage of hrHPV positive women. Among hrHPV positives our study yielded a median M-index value of 8 (15.1) *vs.* 164.8 (573.3) in negative (<CIN2) and positive (CIN2+) confirmed lesions, respectively (Fig. 2) with 35.4 (12.1) and 36.4 (9) median age levels. Based on the above findings, it can be concluded that the significant elevation in POU4F3 methylation reflected the underlying high-grade CIN.

In a recent analysis of the clinical performance POU4F3 was found to be 74% sensitive and 89% specific when used in triage of hrHPV positives by Pun *et al.*[28] Even though DNA methylation provides a quantitative result, currently no generally accepted methods exist to determine the cut-off value for methylation biomarkers. However, the importance to set the qMSP thresholds in order to avoid missing relevant high-grade CIN lesions was previously mentioned.[27] Since the methylation level may show age-related distribution,[46] the optimal cut-off value has to be selected with respect to the age of the subject. Since the median M-index level of CIN2+ cases in the population under 30 years differs significantly from that of the population over 30 years, showing remarkable differences in the distribution of M-index values as well, namely 35.2 (IQR 32.5; min-max 5.9–536.1) *vs.* 255.1 (IQR 864.5; min-max 4.1–5276.2) (Table 1a), it was deemed reasonable to adjust the cut-off threshold respecting these age groups. Consequently, the cut-off value of POU4F3 methylation biomarker test needs to be different in the age group under 30 in order to avoid missing relevant high-grade CIN lesions and achieve high sensitivity.

The PALMS study is one of the largest studies assessing a cervical nonmethylation biomarker, where p16/Ki67 showed 93.3 and 87.8% sensitivity for CIN2+ in women under and over 30 years of age, respectively.[36] In light of these findings, in the TRACE study the cut-off threshold of M-index was determined to be the 93.3th and 87.8th percentile ranking for CIN2+ cases under and over 30 years of age, respectively. Using this threshold, the M-index cut-off level among hrHPV positives was calculated to be 11.3 for subjects aged under 30 and 29.2 for those aged 30 and over, respectively. The overall rate of methylation positivity was 26.6% in the age group of over 25 years ($n = 1,431$, Table 1a).

The characteristics of methylation is presented in Figure 3 by ROC curves with respect to the population under and over 30 years of age. Figure 3 demonstrates the performance of the cut-off value which was determined to be the predefined percentiles ranking for CIN2+ cases in the whole enrolled population under and over 30 years of age.

Evaluation of POU4F3 methylation in HPV triage

As mentioned before, there are differences in the sample characteristics of samples collected from the outpatient clinics and the oncology center population. In the latter one, HPV

Table 1. Study characteristics, women aged 25–65 (A–C). CIN2+ included cervical intraepithelial neoplasia grade 2 (CIN2), cervical intraepithelial neoplasia grade 3 (CIN3), cervical *in situ* carcinoma (CIS), invasive cervical carcinoma (CC), adenocarcinoma *in situ* (AIS), adenoid cystic carcinoma (ACC), respectively. Categorical variables are presented as count (percentage), continuous variables are presented as mean (±SD), median (interquartile range [IQR]), [min]-[max]. HPV test performance was described by sensitivity (SE), specificity (SP), positive predictive value (PPV), negative predictive value (NPV) for endpoints (CIN2+ and CIN3+). Agreement of the HPV tests compared was demonstrated by overall agreement, positive percent agreement (PPA), Chamberlain's PPA, negative percent agreement (NPA) Chamberlain's NPA, Cohen's kappa, prevalence-adjusted and bias-adjusted kappa (PABAK), sensitivity and specificity for cobas® as comparator test (SE and SP for REF).

A. Valid test results of samples collected from women aged 25–65 (n = 5,384)							
Age groups		**25–29**	**30–39**	**40–49**	**50–59**	**60–65**	**Total 25–65**
All samples	LBC sample	1,136 (21.1%)	2,379 (44.2%)	1,269 (23.6%)	458 (8.5%)	142 (2.6%)	5,384 (100%)
hrHPV detection[1]	hrHPV−	714 (17.4%)	1,817 (44.3%)	1,057 (25.8%)	386 (9.4%)	123 (3%)	4,097 (100%)
	hrHPV+	422 (32.8%)	562 (43.7%)	212 (16.5%)	72 (5.6%)	19 (1.5%)	1,287 (100%)
	HPV16	123 (32.7%)	170 (45.2%)	55 (14.6%)	23 (6.1%)	5 (1.3%)	376 (100%)
	HPV18	45 (41.3%)	38 (34.9%)	17 (15.6%)	7 (6.4%)	2 (1.8%)	109 (100%)
	Other HR	339 (34.1%)	426 (42.9%)	164 (16.5%)	50 (5%)	14 (1.4%)	993 (100%)
	Multiplex HPV	82 (44.1%)	70 (37.6%)	24 (12.9%)	8 (4.3%)	2 (1.1%)	186 (100%)
Cytology[2]	NILM	949 (20.5%)	2,035 (44%)	1,086 (23.5%)	424 (9.2%)	134 (2.9%)	4,628 (100%)
	ASCUS	7 (13.7%)	28 (54.9%)	14 (27.5%)	2 (3.9%)	0 (0%)	51 (100%)
	LSIL	165 (26.9%)	268 (43.6%)	146 (23.8%)	29 (4.7%)	6 (1%)	614 (100%)
	ASCH	2 (8.3%)	14 (58.3%)	5 (20.8%)	2 (8.3%)	1 (4.2%)	24 (100%)
	HSIL	13 (20%)	34 (52.3%)	17 (26.2%)	1 (1.5%)	0 (0%)	65 (100%)
	CC	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)
	AGC	0 (0%)	0 (0%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)
M-index distribution in NILMs[3]	Median (IQR)	8 (12.1)	9.8 (16.3)	14 (22.3)	20.6 (33.4)	21.5 (28.8)	11.3 (18.6)
	min-max	0–492.1	0–3449.4	0–1484.2	0.1–775.8	1–4629.4	0–4629.4
M-index distribution in CIN2+ cases[3]	Median (IQR)	35.2 (32.5)	294.3 (879.5)				164.8 (573.3)
			164.7 (392.6)	766.8 (2059.4)	488.4 (1499.3)	1990.8 (550.4)	
	min-max	5.9–536.1	4.1–5276.2				4.1–5276.2
			4.1–5276.2	6.4–2916.3	11.1–3009.7	1440.4–2541.2	
POU4F3 Methylation[4]	MET−	703 (17.8%)	1,940 (49.1%)	928 (23.5%)	294 (7.4%)	88 (2.2%)	3,953 (100%)
	MET+	433 (30.3%)	439 (30.7%)	341 (23.8%)	164 (11.5%)	54 (3.8%)	1,431 (100%)
Confirmation[5]	NEG by second visit	10 (35.7%)	11 (39.3%)	4 (14.3%)	3 (10.7%)	0 (0%)	28 (100%)
	NEG by histology	7 (25%)	10 (35.7%)	6 (21.4%)	3 (10.7%)	2 (7.1%)	28 (100%)
	CIN1	3 (18.8%)	11 (68.8%)	2 (12.5%)	0 (0%)	0 (0%)	16 (100%)
	CIN2	4 (22.2%)	8 (44.4%)	6 (33.3%)	0 (0%)	0 (0%)	18 (100%)
	CIN3	12 (18.5%)	39 (60%)	11 (16.9%)	3 (4.6%)	0 (0%)	65 (100%)
	CIS, CC, AIS, ACC	0 (0%)	5 (41.7%)	5 (41.7%)	0 (0%)	2 (16.7%)	12 (100%)
	<CIN2	20 (27.8%)	32 (44.4%)	12 (16.7%)	6 (8.3%)	2 (2.8%)	72 (100%)
	CIN2+	16 (16.8%)	52 (54.7%)	22 (23.2%)	3 (3.2%)	2 (2.1%)	95 (100%)
	CIN3+	12 (15.6%)	44 (57.1%)	16 (20.8%)	3 (3.9%)	2 (2.6%)	77 (100%)
	Not confirmed	1,100 (21.1%)	2,295 (44%)	1,235 (23.7%)	449 (8.6%)	138 (2.6%)	5,217 (100%)

Table 1. Study characteristics, women aged 25–65 (A–C). (Continued)

B. Test results of carcinoma and carcinoma *in situ* cases ($n = 12$)					
No.	**hrHPV[1] (CONFIDENCE™ & cobas®)**	**Citology[2]**	**M-index[3]**	**Histology[5]**	**Age**
#1	HPV16+	Negative	652	ACC	33
#2	negative	Negative	42	CC	42
#3	HPV16+	ASCH	2,541	CIS	62
#4	HPV18+, other hrHPV+	HSIL	2,227	CC	48
#5	HPV18+	HSIL	2,916	AIS, CIS, CIN3	42
#6	HPV16+	Negative	443	CIS	31
#7	other hrHPV+	Negative	2,929	CIS	36
#8	other hrHPV+	ASCH	2,898	CIS	43
#9	HPV16+, other hrHPV+	HSIL	80	CIS	31
#10	HPV16+	HSIL, CC	1,440	CIS	60
#11	other hrHPV+	HSIL	1,137	CIS	43
#12	HPV16+, HPV18+, other hrHPV+	HSIL	1,768	CIS	36

C. CONFIDENCE HPV™ test results compared to cobas® test ($n = 3,150$).				
25–65, CIN2+	**CONFIDENCE HPV™**	**cobas® HPV**	**Relative ratio**	
SE	95.2 (88.1–98.7)	96.4 (89.8–99.3)	0.99 (0.93–1.05)	
SP	77.8 (76.2–79.2	79.9 (78.5–81.4)	0.97 (0.95–0.99)	
PPV	10.4 (8.2–12.8)	11.5 (9.2–14.1)	0.90 (0.67–1.21)	
NPV	99.8 (99.6–99.9)	99.8 (99.6–99.9)	0.99 (0.99–1.01)	
25–65, CIN3+	**CONFIDENCE HPV™**	**cobas® HPV**	**Relative ratio**	
SE	98.5 (91.8–99.9)	98.5 (91.8–99.9)	1.00 (0.96–1.04)	
SP	77.4 (75.9–78.9)	79.6 (78.1–81.4)	0.97 (0.95–0.99)	
PPV	8.5 (6.7–10.8)	9.4 (7.3–11.8)	0.91 (0.66–1.27)	
NPV	99.9 (99.8–99.9)	99.9 (99.8–99.9)	0.99 (0.99–1.01)	
Agreement	**hrHPV (any)**	**HPV16**	**HPV18**	**Other hrHPV**
Overall agreement	92.3	97.5	99.3	93.4
PPA (cPPA)	83.3 (81.2–85.4) [71.4]	81.1 (77.1–85.2) [68.3]	84.5 (78.1–90.9) [73.2]	81.3 (78.8–83.8) [68.5]
NPA (cNPA)	95.0 (94.4–95.6) [90.5]	98.7 (98.4–98.9) [97.4]	99.6 (99.5–99.8) [99.3]	95.9 (95.4–96.5) [92.3]
Cohen's kappa	0.78 (0.75–0.82)	0.79 (0.76–0.83)	0.84 (0.81–0.88)	0.77 (0.74–0.81)
PABAK	0.85 (0.83–0.86)	0.95 (0.94–0.96)	0.99 (0.98–0.99)	0.87 (0.85–0.88)
SE for REF	87.3 (84.6–89.7)	83.7 (77.9–88.5)	88.2 (78.1–94.8)	85.8 (82.6–88.7)
SP for REF	93.7 (92.7–94.6)	98.4 (97.9–98.9)	99.5 (99.2–99.8)	94.9 (93.9–95.7)

Valid test results of the enrolled samples: [1]CONFIDENCE HPV™; [2]LBC cytology; [3,4]CONFIDENCE Marker™; [5]Confirmation (LBC re-testing in second visit or histology by cone/punch biopsy).

positivity (87%) and the rate of the confirmed cases CIN2+ (99%) were higher, since all of these women were referred for treatment and the negativity rate of the samples collected from these subjects was very low.

Notwithstanding these differences, the medical decisions were based on the same principles and methods both in the outpatient and in the oncology center population. Also the methylation level of CIN2+ samples was found to be similar in the two populations (median M-index level of 158 and 185) (Supporting Information Table S3).

The oncology center samples provided an opportunity to increase the number of confirmed cases used for the verification bias adjustment model allowing to assess the difference between the triage tests compared in the study with higher reliability. Pooling together these samples did not produce any additional bias because the two tests, LBC cytology and methylation achieved comparable crude sensitivity in the outpatient and in the oncology center populations. In addition, both tests found a similar number of false positives among the confirmed negatives of the oncology center population;

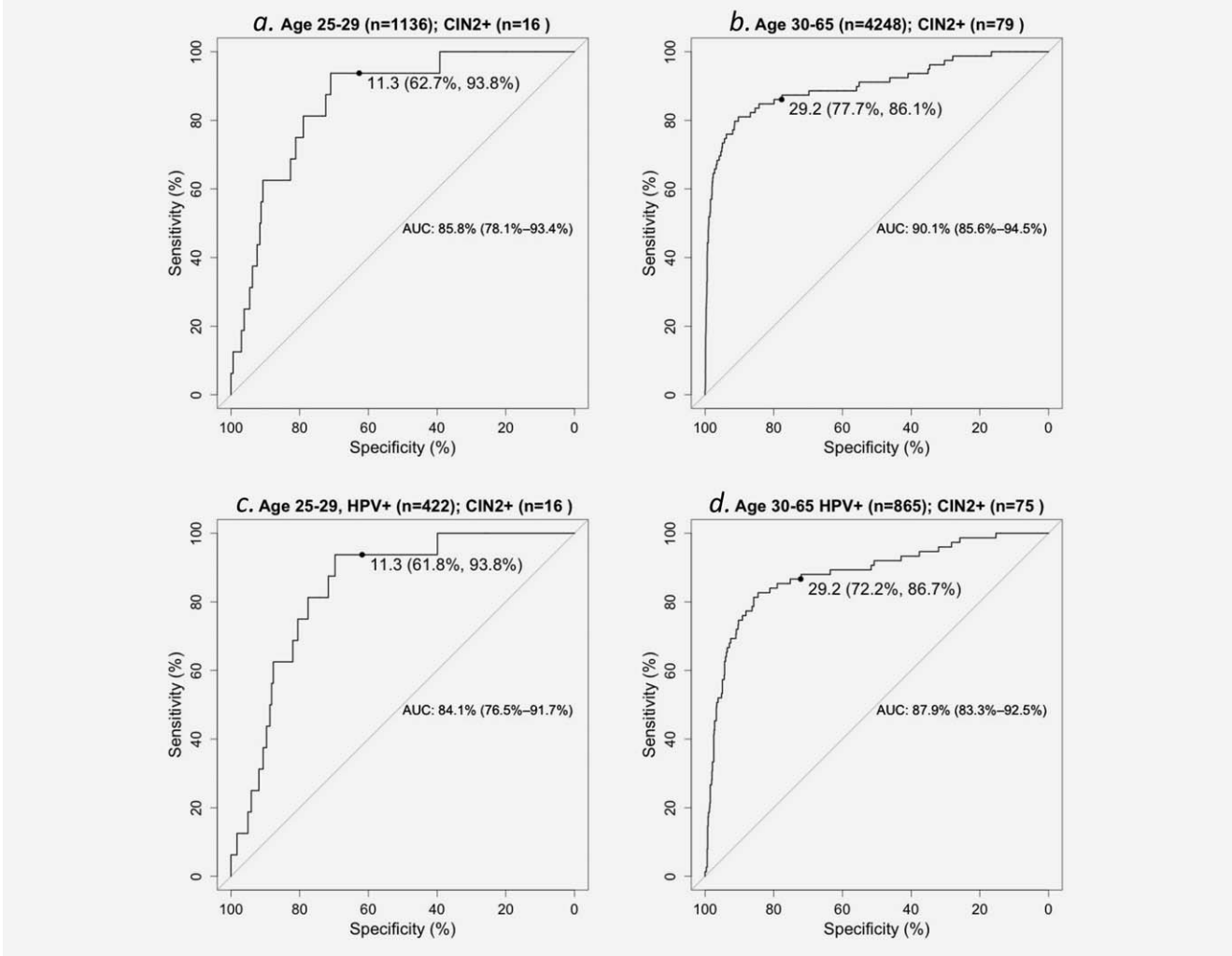

Figure 3. ROC curves (*a–d*). Receiver operating characteristic (ROC) analysis was performed to show the sensitivity and specificity for CIN2+ detection in four subpopulations (age groups between 25–29 and 30–65 years, respectively; hrHPV positives and all patients irrespective of HPV status). The figure demonstrates the performance of the cut-off value which was determined to be the 93.3th and 87.8th percentile ranking for CIN2+ cases in the whole enrolled population (*i.e.*, aged 18–65) under and over 30 years of age (calculated to be 11.3 and 29.2, respectively).

therefore, there was no discrimination of either tests with regard to specificity. Therefore, it was deemed appropriate to evaluate all of the enrolled samples collectively for the relative analysis of different triage strategies as it is provided below.

The clinical performance of POU4F3 methylation was evaluated by comparing the performance of methylation and cytology in triage[6] of hrHPV positives ($n = 1{,}287$). Methylation and cytology combined with HPV16/18 genotyping[5] were also evaluated. Absolute and relative sensitivities (both crude and adjusted) calculated using bootstrap approach are shown in Table 2. The result of cross validation corroborates these findings (Supporting Information Table S5). Both crude and adjusted estimations calculated for CIN2+ and CIN3+ endpoints in the age groups of 25–60, 25–29 and 30–65 were provided (Table 2a, 2b and 2c). All results described below were subjected to verification bias adjustment, if not noted otherwise.

In the hrHPV positive population aged 25–65, sensitivity and specificity of cytology using ASCUS or worse cutoff for CIN2+ was 42.7% (95% CI: 33–56.3) and 80.4% (95% CI: 75.4–85.5), respectively. The adjusted sensitivity and specificity of POU4F3 methylation in discriminating CIN2+ among hrHPV positives aged 25–65 was 70.1% (95% CI: 55.5–84.5) and 81.4% (95% CI: 58.0–95.4), respectively. When comparing POU4F3 and cytology in triage, methylation was found to be significantly more sensitive than cytology: relative sensitivity and specificity yielded 1.67 (95% CI: 1.23–2.24) and 1.01 (95% CI: 0.7–1.22), respectively, in the age group of 25–65 for CIN2+. For CIN3+ endpoint in the age group of 25–65 and 30–65, relative sensitivities were found to be 1.74 (95% CI: 1.25–2.33) and 1.64 (95% CI:

Table 2. Clinical performance of different triage strategies evaluated in the TRACE study population and calculated by bootstrap approach ($R = 1000$) (A–C). Crude and verification bias-adjusted results of different triage strategies: CONFIDENCE Marker™ compared to LBC cytology and both combined with genotyping in the triage of hrHPV positives (CONFIDENCE HPV™). Point estimates of absolute sensitivities (SE), absolute specificities (SP), and relative sensitivity and specificity values are described with 95% confidence intervals (95%CI). Bootstrap approach (number of replicates = 1,000) was used to calculate the internally validated point estimates and confidence intervals.

A. Study population aged 25–65 (total $n = 5{,}384$; hrHPV+ $n = 1.287$)									
		CIN2+					CIN3+		
		POU4F3 marker	LBC cytology	Relative ratio			POU4F3 marker	LBC cytology	Relative ratio
Crude	SE	88.2 (85.5–91.0)	80.1 (71.6–88.2)	1.10 (1.00–1.24)	Crude	SE	89.6 (84.6–94.6)	80.1 (70.9–88.7)	1.12 (0.99–1.27)
	SP	72.9 (50.8–89.5)	54.7 (41.4–68.3)	1.36 (0.86–1.95)		SP	60.9 (41.0–77.9)	47.4 (36.1–59.3)	1.31 (0.82–1.87)
Adjusted	SE	70.1 (55.5–84.5)	42.7 (33.0–56.3)	1.67 (1.23–2.24)	Adjusted	SE	72.3 (58.5–86.8)	42.3 (31.2–55.7)	1.74 (1.25–2.33)
	SP	81.4 (58.0–95.4)	80.4 (75.1–85.5)	1.01 (0.70–1.22)		SP	76.5 (52.1–91.9)	77.8 (73.0–82.9)	0.98 (0.66–1.21)
		POU4F3 marker with 16/18 typing	LBC cytology with 16/18 typing	Relative ratio			POU4F3 marker with 16/18 typing	LBC cytology with 16/18 typing	Relative ratio
Crude	SE	92.3 (86.3–97.5)	93.3 (87.4–98.0)	0.98 (0.92–1.06)	Crude	SE	93.4 (87.3–98.4)	93.3 (87.1–98.5)	1.01 (0.94–1.08)
	SP	43.2 (30.4–57.4)	28.2 (16.3–40.4)	1.59 (1.00–2.56)		SP	37.1 (25.3–48.7)	23.6 (13.5–34.3)	1.64 (1.06–2.70)
Adjusted	SE	82.3 (69.9–95.1)	75.3 (62.2–89.7)	1.09 (0.92–1.28)	Adjusted	SE	82.7 (67.2–96.6)	74.8 (59.6–90.2)	1.11 (0.94–1.31)
	SP	67.5 (58.0–76.8)	65.0 (56.5–72.2)	1.04 (0.85–1.23)		SP	61.4 (52.3–69.3)	59.5 (50.5–66.8)	(0.88–1.21)
B. Study population aged 25–29 (total $n = 1{,}136$; hrHPV+ $n = 422$)									
		CIN2+					CIN3+		
		POU4F3 marker	LBC cytology	Relative ratio			POU4F3 marker	LBC cytology	Relative ratio
Crude	SE	92.2 (78.6–100.0)	56.1 (29.4–81.3)	1.75 (1.13–3.00)	Crude	SE	89.6 (71.4–100.0)	58.4 (26.7–85.7)	1.67 (1.00–3.49)
	SP	71.9 (31.9–100.0)	83.9 (40.0–86.7)	1.18 (0.47–2.20)		SP	59.6 (27.3–85.2)	61.2 (40.0–80.9)	1.01 (0.44–1.70)
Adjusted	SE	83.2 (57.5–100.0)	31.9 (18.9–52.5)	2.76 (1.55–4.46)	Adjusted	SE	78.6 (44.4–100.0)	34.4 (15.1–65.8)	2.57 (1.00–5.19)
	SP	79.2 (51.6–100.0)	74.1 (64.9–82.6)	1.08 (0.66–1.44)		SP	70.2 (45.6–89.6)	74.1 (65.7–80.9)	0.95 (0.62–1.26)
		POU4F3 marker with 16/18 typing	LBC cytology with 16/18 typing	Relative ratio			POU4F3 marker with 16/18 typing	LBC cytology with 16/18 typing	Relative ratio
Crude	SE	93.5 (76.9–100.0)	87.3 (66.7–100.0)	1.08 (1.00–1.30)	Crude	SE	91.3 (71.4–100.0)	91.3 (71.4–100.0)	1.00
	SP	52.7 (30.8–75.0)	36.8 (14.3–60.0)	1.57 (0.86–3.00)		SP	43.4 (22.7–66.6)	34.3 (13.3–56.0)	1.37 (0.7–2.67)
Adjusted	SE	89.5 (67.7–100.0)	72.7 (45.1–100.0)	1.28 (1.00–1.86)	Adjusted	SE	84.5 (55.1–100.0)	84.5 (55.1–100.0)	1.00
	SP	62.0 (45.3–81.4)	54.6 (38.3–69.7)	1.17 (0.77–1.66)		SP	50.6 (37.1–69.7)	54.5 (36.1–67.8)	(0.72–1.60)

Table 2. Clinical performance of different triage strategies evaluated in the TRACE study population and calculated by bootstrap approach ($R = 1000$) (A–C). (Continued)

C. Study population aged 30–65 (total $n = 4,248$; hrHPV+ $n = 865$)									
		CIN2+					CIN3+		
		POU4F3 marker	LBC cytology	Relative ratio			POU4F3 marker	LBC cytology	Relative ratio
Crude	SE	87.3 (84.3–90.0)	85.3 (76.9–93.0)	1.03 (0.94–1.15)	Crude	SE	89.6 (84.0–95.1)	84.3 (75.0–92.5)	1.07 (0.95–1.21)
	SP	74.4 (48.0–93.9)	49.9 (32.5–65.6)	1.54 (0.88–2.40)		SP	62.4 (38.1–84.1)	40.1 (25.0–53.8)	1.61 (0.87–2.64)
Adjusted	SE	65.6 (47.8–83.2)	46.2 (33.2–66.3)	1.47 (0.94–2.16)	Adjusted	SE	71.4 (52.8–90.2)	44.7 (30.5–65.2)	1.64 (1.08–2.27)
	SP	82.2 (45.8–98.3)	82.2 (75.3–88.2)	1.00 (0.54–1.25)		SP	79.7 (44.0–97.2)	78.7 (72.4–84.8)	1.02 (0.54–1.28)
		POU4F3 marker with 16/18 typing	LBC cytology with 16/18 typing	Relative ratio			POU4F3 marker with 16/18 typing	LBC cytology with 16/18 typing	Relative ratio
Crude	SE	91.9 (84.8–97.4)	94.6 (89.0–98.8)	0.97 (0.89–1.05)	Crude	SE	93.8 (87.3–98.6)	93.8 (86.9–98.7)	1.00 (0.92–1.09)
	SP	38.4 (22.9–55.6)	23.4 (10.0–38.5)	1.79 (0.91–3.67)		SP	33.2 (20.0–46.8)	17.8 (7.1–29.3)	2.09 (1.00–4.50)
Adjusted	SE	80.2 (60.3–94.7)	77.4 (58.3–95.0)	1.04 (0.89–1.22)	Adjusted	SE	81.8 (60.82–97.5)	74.5 (54.4–94.6)	1.11 (0.92–1.33)
	SP	70.4 (55.8–95.0)	67.7 (52.3–77.7)	1.05 (0.87–1.28)		SP	65.6 (52.1–75.8)	60.2 (46.0–70.2)	1.10 (0.91–1.37)

Figure 4. Number of cases identified by the test results among hrHPV positives. Distribution of positives into different categories as per POU4F3 methylation positivity (455 MET+) and abnormal LBC (364 ASCUS+) results. Number of cases confirmed as CIN2+ (91) and <CIN2 (53) are also shown. HPV was detected by CONFIDENCE HPV™ test.

1.08–2.27), respectively. Based on the above, it is reasonably assumed that methylation of POU4F3 can achieve significantly higher sensitivity than cytology-based triage with comparable specificity (Table 2).

Sensitivity increased, whereas specificity decreased when cytology and methylation were used in combination with HPV16/18 genotyping. In case of combined triage of cytology and genotyping, 75.3% (95% CI: 62.2–89.7) sensitivity and 65.0% (95% CI: 56.5–72.2) specificity were estimated for CIN2+ (Table 2). POU4F3 marker combined with genotyping showed 82.3% (95% CI: 69.9–95.1) sensitivity and 67.5% (95% CI: 58.0–76.8) specificity for CIN2+ in the population aged 25–65. When comparing cytology and POU4F3 methylation both combined with HPV16/18 genotyping, methylation-based triage was found to be significantly more specific than cytology based triage with comparable sensitivity for CIN2+ in crude analysis, *i.e.*, crude relative specificity was 1.59 (95% CI: 1.00–2.56) and relative sensitivity was 0.98 (95% CI: 0.92–1.06) in the age group of 25–65, respectively. The performance of genotyping combined triage methods was similar in verification bias-adjusted analysis. The triage strategies under evaluation showed similar differences for CIN3+ and in the age group of 30–65, respectively (Table 2).

To assess whether the HPV test used had an impact on the performance, a separate *post hoc* analysis and comparison were implemented on a subset of samples collected from women aged 25–65 (total of $n = 3,150$): first, the CONFIDENCE Marker™ applied with the CONFIDENCE HPV™ test ($n = 761$ HPV+) and second, the CONFIDENCE

Marker™ applied with the cobas® HPV test ($n = 694$ HPV+). The clinical performance was found to be independent of the type of HPV test used (Supporting Information Table S7).

Distribution of referred number of CIN2+ cases

Out of the total of 1,287 hrHPV positive women over 25 years of age, 364 ASCUS+ and 455 MET+ (methylation positive) cases were detected. From the total 364 ASCUS+ cases, 97 were referred for confirmation: 73 and 24 were verified as CIN2+ and <CIN2, respectively.

Out of the 91 histologically confirmed CIN2+ patients, overall 66 would have been referred based on the result of the methylation test, which is 90% of the 73 ASCUS+ and CIN2+ cases. Consequently, POU4F3 was found to be a remarkable marker with excellent potency to identify the severe cases. Eighty-eight further samples with increased methylation level and ASCUS positivity may be verfied by the ongoing follow-up phase of the study (Fig. 4).

The distribution of confirmed CIN2+ and <CIN2 cases among ASCUS+ and MET+ samples reflects the performance of the two tests in triage. Among hrHPV positives 7 *vs.* 14 of the CIN2+ cases were referred based on cytology results alone and would have been referred based on methylation results, respectively. However, cytology and methylation alone identified 19 *vs.* 10 false positive samples from <CIN2 confirmed cases (Fig. 4).

These differences among cases with verified disease can be explained by the lower sensitivity and specificity of cytology compared to POU4F3 methylation among hrHPV positive confirmed cases. The results of the ongoing follow-up phase may provide further opportunities for a more accurate evaluation of the clinical performance in different triage strategies.

Discussion

The goal of the ongoing TRACE trial was to demonstrate the clinical utility of POU4F3 as a biomarker for cervical lesions by *post hoc* analysis and the study aimed to clinically validate the CONFIDENCE™ assay. The trial involves >6,000 women, which makes it one of the largest epigenetic biomarker validation studies conducted up to date.[25]

The CONFIDENCE HPV™ test showed good agreement in comparison with the cobas® HPV test considering the results of the HORIZON study where the performance of four HPV tests were evaluated (HC2, cobas, CLART, APTIMA), and 42–58% positive agreement was found on any compared pair of assays with 0.53–0.75 kappa value.[47]

In the TRACE study population, the promoter sequence methylation of the POU4F3 gene showed significant elevation in high-grade CIN cases independently of the hrHPV status. The median M-index was 164.8 (573.3) *vs.* 8 (15.1) in CIN2+ cases and in samples with no high-grade lesions, respectively. This significant difference in the M-index values highlight the discriminative power of the quantitative nature of the POU4F3 biomarker.

As the current analysis demonstrated, the performance of the POU4F3 biomarker was characterized by high sensitivity. The advantage of POU4F3 marker over cytology was shown in the triage of women with hrHPV infection. The relative sensitivity of POU4F3 was found to be 1.74 (95% CI: 1.25–2.33) and 1.64 (95% CI: 1.08–2.27) for CIN3+ endpoint in the age group of 25–65 and 30–65, respectively, without a compromise on specificity. As mostly test positive subjects were referred for verification, we used weighted generalized estimating equations to adjust for this verification bias.[44] The application of bootstrap approach and repeated k-fold cross-validation offered a rigorous internal validation in the analytical step, thereby improving the general applicability of the results. These findings provided solid evidence that molecular cervical cancer screening is feasible using hrHPV as first line method and POU4F3 methylation as triage method.

The study has limitations which need to be pointed out. The performance of the tests in the outpatient and oncology center study populations may not be fully applicable to a screening population. Colposcopy centers are not part of the Hungarian screening protocol; nevertheless, all women underwent colposcopy-assisted VIA in the study. Due to the conditions set in the Ethics Committee approval, it was not possible to alter the management of the patients for the purpose of the study, including referring patients with negative test results or based on the methylation level in their sample for verification. Our verification bias-adjusted analysis lived on the assumption that the test results were the only variables that influenced whether or not diagnostic verification was performed.

The study samples were tested for POU4F3 methylation only once without replicates; however, the reproducibility of the assay was determined satisfactory based on control DNA tests and a subset analysis of clinical samples.

Given that an age dependent shift of the median M-index level was observed from 25 to 65 years of age, a complex age-related cut-off optimization may be beneficial to further improve the study outcomes and the clinical performance of POU4F3. A revision of the ethical approval will make it possible to follow-up and confirm cases indicated also by methylation positivity.

In conclusion, POU4F3 alone showed significantly higher sensitivity than LBC-based triage with comparable specificity. The ongoing follow-up phase of the study may support the verification bias-adjusted estimations presented above and may lead to a more accurate evaluation of the clinical performance of POU4F3 DNA methylation. On the basis of our findings and further literature data,[28] the promoter sequence methylation of POU4F3 gene as a triage test for hrHPV positives appears to be a noteworthy method. We can reasonably assume that its quantitative nature offers the potential for a more objective and discriminative risk assessment tool[16,48,49] in the prevention and diagnostics of high-grade CIN lesions and cervical cancer.

Acknowledgements

The authors wish to extend their thanks to all of the clinicians whose professionalism was substantial in the care of patients enrolled in the study:

Gusztáv Adorján[1], Edit Babarczi[2], József Benedek[1], Klára Boiskó[3], János Gál[1], Ádám Galamb[4], Péter Hazslinszky[2], Dániel Horányi[5], Béla Ivancsó[1], János Kornafeld[1], Árpád Kosztin[5], Róbert Kovács[1], Gábor Lázár[5], Tibor Leskó[1], János Losonczi[1], Róbert Lukács[1], László Medve[1], Kinga Oláh[3], László Pálfalvi[5], Judit Péli-Tóth[1], Ottó Petróczy[1], Zoltán Pósz[1], Krisztina Rácz[1], Pál Siklós[5], Péter Sziklai[5], Nóra Teiringer[1], Hajnalka Varsányi[5], and Eszter Vonnák[4].

We gratefully acknowledge the following individuals whose assistance was essential for the initiation of the project: Fruzsina Ancza[6,7], Tiborné Balogh[1], Marcella Dombrádi[5,8], Evelin Farkas-Volf[6,7], Anett Kovács[6,7], Laura Kovács[7], Zsuzsanna Liszonyi[9], Johanna Mózes[7], Zsuzsa Ottó[10], Magdolna Pekár[4,11], Mátyás Péter[7], Magdolna Szabó[9], Erzsébet Szarka[6,7], Ágnes Tényiné Gót[3], and Norbert Varga[7]. Their contributions are much appreciated.

[1]Jósa András Regional Central Hospital, Nyíregyháza, Hungary; [2]Department of Pathology, St. Stephan Hospital, Budapest, Hungary; [3]Healthcare Center of Ferencváros, Budapest, Hungary; [4]Amora Kft., Budapest, Hungary; [5]Department of Obstetrics and Gynaecology, St. Stephan Hospital, Budapest, Hungary; [6]NEUMANN Diagnostics Ltd., Budapest, Hungary; [7]Cellcall Ltd., Budapest, Hungary; [8]Profilakt Ltd., Budapest, Hungary; [9]SYNLAB Hungary Ltd. GenoID Molecular Diagnostic Laboratory, Budapest, Hungary; [10]MEDSERV Ltd., Budapest Hungary; [11]Second Department of Pathology, Semmelweis University, Budapest, Hungary.

The study was approved by the Review Board of the Hungarian National Scientific and Ethical Committee (ETT TUKEB, Hungary, approval number: V-R-021/04346–4/2013) for clinical and laboratory sites.

The study was conducted in line with the International Conference on Harmonization Guideline for Good Clinical Practice and written informed consent was obtained from all patients involved.

Conflict of Interest

A. Kocsis is an employee of Cellcall and NEUMANN. T. Takács is a part time employee of Cellcall, he receives consultancy fees from NEUMANN, he is minority owner of NEUMANN, the inventor of the HPV detection technology described in the manuscript covering which Cellcall submitted a patent application (license now owned by NEUMANN Diagnostics Kft.); he is a part time employee of SYNLAB GenoID Laboratory. Cs. Jeney was a former employee of Cellcall. T. Ferenci receives consultancy fees from NEUMANN. HC Lai is the inventor of patent application covering the methylation testing of the POU4F3 gene which was licensed from Prof Lai and the National Defense Medical Center of Taiwan. B. Jaray and Zs. Schaff have no conflict of interest. There is no conflict of interest to report in case of R. Koiss, G. Sobel, K. Pap and I. Szekely. We would like to mention nevertheless that the doctors and clinicial teams received clinical investigator fees covering related expenses. M. Nyíri is an employee of Cellcall and NEUMANN. M. Benczik receives consultancy fees from Cellcall and NEUMANN, she is an employee of SYNLAB GenoID Laboratory, which participated in the TRACE clinical study (all related expenses were charged to Cellcall and NEUMANN).

References

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2012;49:1374–403.
2. Wright TC Jr. Cervical cancer screening in the 21st century: is it time to retire the PAP smear?. *Clin Obstet Gynecol* 2007;50: 313–23.
3. Cuzick J, Clavel C, Petry KU, et al. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. *Int J Cancer* 2006;119:1095–101.
4. Sahasrabuddhe VV, Luhn P, Wentzensen N. Human papillomavirus and cervical cancer: biomarkers for improved prevention efforts. *Future Microbiol* 2011;6:1083–98.
5. Huh WK, Ault KA, Chelmow D, et al. Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance. *Gynecol Oncol* 2015;136:178–82.
6. Karsa L, Arbyn M, Vuyst H. European guidelines for quality assurance in cervical cancer screening. Summary of the supplements on HPV screening and vaccination. *Papillomavirus Res* 2015;1:22–31.
7. 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. *Am J Clin Pathol* 2012;137:516–42.
8. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010;11:249–57.
9. Anttila A, Kotaniemi-Talonen L, Leinonen M, et al. Rate of cervical cancer, severe intraepithelial neoplasia, and adenocarcinoma in situ in primary HPV DNA screening with cytology triage: randomised study within organised screening programme. *Br Med J* 2010;340:c1804.
10. Gage JC, Schiffman M, Katki HA, et al. Reassurance against future risk of precancer and cancer conferred by a negative human papillomavirus test. *J Natl Cancer Inst* 2014;106:dju153.
11. Ronco G, Dillner J, Elfstrom KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet* 2014;383: 524–32.
12. 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. *Br Med J* 2008; 337:a1754.
13. Rijkaart DC, Berkhof J, van Kemenade FJ, et al. HPV DNA testing in population-based cervical screening (VUSA-Screen study): results and implications. *Br J Cancer* 2012;106:975–81.
14. Kitchener HC, Gilham C, Sargent A, et al. A comparison of HPV DNA testing and liquid based cytology over three rounds of primary cervical screening: extended follow up in the ARTISTIC trial. *Eur J Cancer* 2011;47:864–71.
15. Tornesello ML, Buonaguro L, Giorgi-Rossi P, et al. Viral and cellular biomarkers in the diagnosis of cervical intraepithelial neoplasia and cancer. *Biomed Res Int* 2013;2013:519619.
16. Schiffman M, Wentzensen N. A suggested approach to simplify and improve cervical screening in the United States. *J Low Genit Tract Dis* 2016;20:1–7.
17. Wentzensen N, Sherman ME, Schiffman M, et al. Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science. *Gynecol Oncol* 2009;112:293–9.
18. Yang N, Nijhuis ER, Volders HH, et al. Gene promoter methylation patterns throughout the process of cervical carcinogenesis. *Cell Oncol* 2010;32:131–43.
19. De Strooper LM, Meijer CJ, 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.
20. Kim MK, Lee IH, Lee KH, et al. DNA methylation in human papillomavirus-infected cervical cells is elevated in high-grade squamous intraepithelial lesions and cancer. *J Gynecol Oncol* 2016; 27:e14.
21. Lai HC, Lin YW, Huang RL, et al. Quantitative DNA methylation analysis detects cervical intraepithelial neoplasms type 3 and worse. *Cancer* 2010;116:4266–74.
22. Overmeer RM, Henken FE, Bierkens M, et al. Repression of MAL tumour suppressor activity by promoter methylation during cervical carcinogenesis. *J Pathol* 2009;219:327–36.
23. Overmeer RM, Louwers JA, Meijer CJ, et al. Combined CADM1 and MAL promoter methylation analysis to detect (pre-)malignant cervical lesions in high-risk HPV-positive women. *Int J Cancer* 2011;129:2218–25.
24. De Strooper LM, Hesselink AT, Berkhof J, et al. Combined CADM1/MAL methylation and cytology testing for colposcopy triage of high-risk HPV-positive women. *Cancer Epidemiol Biomarkers Prev* 2014;23:1933–7.
25. Verhoef VM, Bosgraaf RP, van Kemenade FJ, et al. Triage by methylation-marker testing versus cytology in women who test HPV-positive on self-collected cervicovaginal specimens (PROHTECT-3): a randomised controlled non-inferiority trial. *Lancet Oncol* 2014;15:315–22.

26. De Strooper LM, van Zummeren M, Steenbergen RD, et al. CADM1, MAL and miR124-2 methylation analysis in cervical scrapes to detect cervical and endometrial cancer. *J Clin Pathol* 2014;67: 1067–71.
27. Hesselink AT, Heideman DA, Steenbergen RD, et al. Combined promoter methylation analysis of CADM1 and MAL: an objective triage tool for high-risk human papillomavirus DNA-positive women. *Clin Cancer Res* 2011;17:2459–65.
28. Pun PB, Liao YP, Su PH, et al. Triage of high-risk human papillomavirus-positive women by methylated POU4F3. *Clin Epigenetics* 2015;7:85.
29. Pepe MS. The statistical evaluation of medical tests for classification and predictioned, vol. xvi. Oxford: Oxford University Press, 2004. 302 p.
30. Arbyn M, Sasieni P, Meijer CJ, et al. Clinical applications of HPV testing: a summary of meta-analyses. *Vaccine* 2006;24(Suppl. 3):S3/78–89.
31. Bosze P, Gocze P, Hernadi Z, et al. The guideline for cervical cancer screening: the Hungarian protocol. A mehnyakrak szuresenek szempontjai: hazai iranyelvek. *Nogyogyaszati Onkol* 2009;14: 11–7.
32. Bornstein J, Bentley J, Bosze P, et al. 2011 colposcopic terminology of the International Federation for cervical pathology and colposcopy. *Obstet Gynecol* 2012;120:166–72.
33. Solomon D, Davey D, Kurman R, MA, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *J Am Med Assoc* 2002;287:2114–9.
34. Huang RL, Chang CC, Su PH, et al. Methylomic analysis identifies frequent DNA methylation of zinc finger protein 582 (ZNF582) in cervical neoplasms. *PLoS One* 2012;7:e41060.
35. Chen YC, Huang RL, Huang YK, et al. Methylomics analysis identifies epigenetically silenced genes and implies an activation of beta-catenin signaling in cervical cancer. *Int J Cancer* 2013; 135:117–27.
36. Ikenberg H, Bergeron C, Schmidt D, et al. Screening for cervical cancer precursors with p16/Ki-67 dual-stained cytology: results of the PALMS study. *J Natl Cancer Inst* 2013;105:1550–7.
37. Wright TC, Stoler MH, Behrens CM, et al. Primary cervical cancer screening with human papillomavirus: end of study results from the ATHENA study using HPV as the first-line screening test. *Gynecol Oncol* 2015;136:189–97.
38. Jeney C, Takacs T, Sebe A, et al. Detection and typing of 46 genital human papillomaviruses by the L1F/L1R primer system based multiplex PCR and hybridization. *J Virol Methods* 2007;140:32–42.
39. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551–8.
40. Graham P, Bull B. Approximate standard errors and confidence intervals for indices of positive and negative agreement. *J Clin Epidemiol* 1998;51:763–71.
41. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993;46:423–9.
42. Efron B, Tibshirani R. An introduction to the bootstrapped, vol. XVI. New York: Chapman & Hall, 1993. 436 p.
43. Geisser S. The predictive sample reuse method with applications. *J Am Stat Assoc* 1975;70:320–8.
44. Xue X, Kim MY, Castle PE, et al. A new method to address verification bias in studies of clinical screening tests: cervical cancer screening assays as an example. *J Clin Epidemiol* 2014;67:343–53.
45. 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: 829–46.
46. Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. *Aging Cell* 2015; 14:924–32.
47. Rebolj M, Preisler S, Ejegod DM, et al. Disagreement between human papillomavirus assays: an unexpected challenge for the choice of an assay in primary cervical screening. *PLoS One* 2014; 9: e86835.
48. Castle PE, Sideri M, Jeronimo J, et al. Risk assessment to guide the prevention of cervical cancer. *Am J Obstet Gynecol* 2007;197:356e1–e6.
49. Wentzensen N. Triage of HPV-positive women in cervical cancer screening. *Lancet Oncol* 2013; 14:107–9.