

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Interpretation of p16INK4a/Ki-67 dual immunostaining for the triage of human papillomavirus-positive women by experts and nonexperts in cervical cytology

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/153438> since

Published version:

DOI:10.1002/cncy.21511

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This is the accepted version of the following article:

Allia, E., Ronco, G., Coccia, A., Luparia, P., Macrì, L., Fiorito, C., Maletta, F., Deambrogio, C., Tunesi, S., De Marco, L., Gillio-Tos, A., Sapino, A. and Ghiringhello, B. (2014), Interpretation of p16INK4a/Ki-67 dual immunostaining for the triage of human papillomavirus-positive women by experts and nonexperts in cervical cytology. *Cancer Cytopathology*. doi: 10.1002/cncy.21511,

which has been published in final form at

<http://onlinelibrary.wiley.com/doi/10.1002/cncy.21511/abstract>

Keywords:

p16INK4a/Ki-67; cervical cancer; human papillomavirus (HPV)-positive women; interpretation; readers

Abstract**BACKGROUND**

The triage of human papillomavirus (HPV)-positive women is needed to avoid overreferral to colposcopy. p16INK4a immunostaining is an efficient triage method. p16INK4a/Ki-67 dual staining was introduced mainly to increase reproducibility and specificity compared with stand-alone p16INK4a staining.

METHODS

Within a pilot project, HPV-positive women were referred to colposcopy if cytology was abnormal or unsatisfactory or HPV testing was still positive after 1 year. For 500 consecutive women, a slide obtained during colposcopy was immunostained for p16INK4a/Ki-67 and independently interpreted by 7 readers without previous experience with dual staining. Four of these readers were experts in cervical pathology and 3 were not. Kappa values for multiple raters, sensitivity, and specificity for cervical intraepithelial neoplasia type 2-positive histology were computed. Because women with normal cytology were underrepresented, estimates for all HPV-positive women were obtained as weighted means of cytology-specific estimates.

RESULTS

The overall kappa for HPV-positive women was 0.70 (95% confidence interval [95% CI], 0.60-0.77). Kappa values were not found to be significantly different between expert and nonexpert readers with regard to cervical cytology but were significantly increased ($P = .0066$) after consensus discussion. The overall specificity estimate for HPV-positive women was 64.0% (95% CI, 57.4%-70.2%): 66.7% (95% CI, 59.8%-73.0%) for experts and 60.5% (95% CI, 59.8%-73.0%) for nonexperts. Among women with abnormal cytology, the sensitivity was 85.5% (95% CI, 77.9%-90.8%): 85.8% (95% CI, 77.9%-91.2%) for experts and 85.1% (95% CI, 76.6%-90.9%) for nonexperts.

CONCLUSIONS

p16INK4a/Ki-67 immunostaining demonstrated good reproducibility and specificity when triaging HPV-positive women. Dual-staining interpretation can be performed, after short training, even by staff who are not experts in cervical cytology. This allows HPV-based screening with triage to be performed in settings in which such expert staff is not available. *Cancer (Cancer Cytopathol)* 2014. © 2014 American Cancer Society.

INTRODUCTION

Cervical screening based on human papillomavirus (HPV) testing has been shown to allow for the earlier detection of persistent high-grade cervical intraepithelial neoplasia (CIN) compared with cytology-based screening,¹⁻⁴ and to be more efficacious in preventing invasive cervical carcinoma.^{5, 6} However, because the majority of HPV infections regress spontaneously without causing high-grade CIN, HPV testing is less specific than cytology, resulting in unacceptably high referral to colposcopy if all HPV-positive women are referred. Therefore, approaches with better specificity, based on further triage of HPV-positive women, are needed.

p16INK4a has an antiproliferative effect during regular cell cycles. In terminally differentiated epithelial cells, p16INK4a expression is downregulated to levels typically not detectable by immunocytochemistry.^{7, 8} Conversely, p16INK4a has been shown to be strongly overexpressed in cervical neoplasia as a consequence of the functional inactivation of the retinoblastoma protein, mediated by the E7 high-risk HPV oncoprotein^{9, 10} that promotes cell cycle deregulation. p16INK4a overexpression has been studied as a method of triaging women with borderline or low-grade cytology and those who are HPV positive.¹¹ HPV-based screening with p16INK4a triage was found to result in increased cross-sectional¹² and longitudinal¹³ sensitivity compared with cytology-based screening, without an increase in referral to colposcopy. However, because p16INK4a can also be overexpressed in normal metaplastic or atrophic cells, morphologic criteria, entailing subjective interpretation, also are needed.^{14, 15} A high variability between studies with regard to the percentage of p16INK4a-positive samples within the same cytological category, suggesting low reproducibility, has been observed.¹⁶

Dual staining with p16INK4a and Ki-67 was introduced to overcome this problem and increase specificity. Ki-67, a proliferation-associated protein, can be detected exclusively in the nucleus of proliferating cells, whereas cells in G₀ phase do not express it.^{17, 18} Because cells that overexpress p16INK4a can actively proliferate only after impairment of the cell-cycle control mechanism, the expression of both Ki-67 and p16INK4a within the same cell mutually exclude each other under normal physiological conditions. Thus, concomitant expression of the 2 proteins in the same cell may be used as an indicator of deregulation of cell cycle control.

In the current study, we examined the reproducibility and accuracy, particularly the specificity, of double staining among HPV-positive women within the framework of a randomized pilot project of HPV-based screening in Turin, Italy, comparing interpretation between experts and nonexperts in cervical cytology.

MATERIALS AND METHODS

Setting, Study Population, and Material

The Turin cervical screening program regularly invites all resident women who are within the target age group. In a pilot project started in March 2010, women aged 35 to 64 years were randomly invited to undergo cervical screening with conventional cytology or HPV testing as the primary screening test. After written informed consent was obtained and approved by the ethical committee (N. CEI/585), women who were invited to the HPV arm underwent an HPV DNA test performed by Hybrid Capture 2 (Qiagen, Hilden, Germany) and were considered to be positive if the relative light unit/cutoff ratio was ≥ 1 , according to the manufacturer's instructions. In the HPV arm, a sample for conventional cytology (Papanicolaou smear) was obtained from all women but was stained and interpreted only if the HPV test was positive. Interpretation was performed centrally according to the 2001 Bethesda System. Each slide was screened by 1 of 5

cytologists and 1 pathologist reviewed the slides considered to be suspicious. If the cytology was determined to be atypical squamous cells of undetermined significance (ASC-US) or more severe or unsatisfactory, women were referred to colposcopy, whereas women who were HPV positive but cytology negative were referred to colposcopy only if an HPV test repeated 1 year later was persistently positive (Fig. 1).

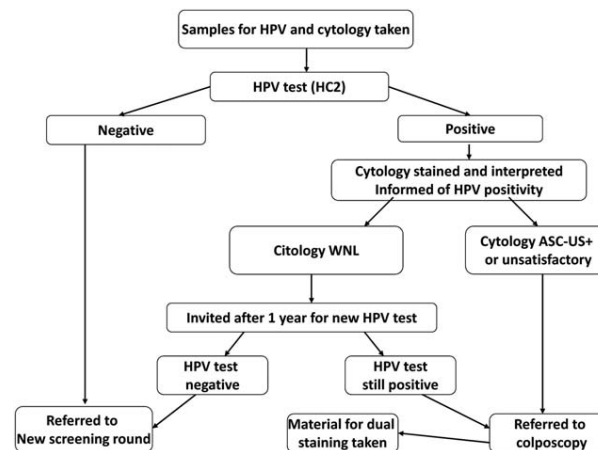


Figure 1. The Turin pilot project is shown demonstrating the protocol in the human papillomavirus (HPV) arm. HC2 indicates Hybrid Capture 2; WNL, within normal limits; ASC-US+, atypical squamous cells of undetermined significance or more severe.

During colposcopies, all suspicious areas were biopsied. All histology was first interpreted by 1 pathologist and reviewed by another with >30 years of experience in cervical pathology (B.G.).

At the time of the first colposcopy, a sample of cervical cells was also routinely taken from each woman and stored in PreservCyt solution (Hologic Inc, Marlborough, Mass) for liquid-based cytology. The current study was conducted on material obtained from the first 500 sequential samples, which were collected between May 2010 and June 2012 (therefore collected after the sample for HPV testing). This material was used to perform as many p16INK4a/Ki-67 tests. Collecting supplementary material along with material for HPV testing from all screened women (the majority of whom were expected to be HPV negative) was not considered to be feasible and destained conventional slides not representative of good-quality dual staining. As a result, women with baseline negative cytology were included only if they had 2 previous positive HPV tests at 1-year intervals. Because this study was performed at the beginning of the pilot project, women with normal cytology were underrepresented compared with their frequency among HPV-positive women.

Immunocytochemical Dual Staining

Slides were prepared from liquid-based cytological samples, using the residual material left after the preparation of 1 slide, on a ThinPrep 2000 Processor (Hologic Inc), fixed in 99% ethanol for 10 minutes, and air-dried overnight. Slides were then immunostained in an automated station (Autostainer Link 48; Dako, Glostrup, Denmark) using the CINtec Plus Cytology kit (Roche Diagnostics, Basel, Switzerland) according to the manufacturer's instructions. Briefly, a cocktail of a monoclonal mouse antibody (clone E6H4) directed toward the human p16INK4a protein and a monoclonal rabbit antibody (clone 274-11 AC3) directed against the human Ki-67 protein were followed by incubation with secondary antibodies conjugated with horseradish peroxidase and with alkaline phosphatase. Horseradish peroxidase-mediated conversion of 3,30-diaminobenzidine chromogen and alkaline phosphatase-mediated conversion of fast red chromogen

led to brown and red staining, respectively, at the p16INK4a and Ki-67 antigen sites. High-grade squamous intraepithelial lesion slides were used as a positive control for each immunoreaction run.

Dual-Staining Interpretation

Each slide was reviewed independently and blindly with respect to cytology and histology once by each of the 7 readers. None of the readers had previous experience with dual staining. Four readers (3 cytotechnologists and 1 pathologist) were experts in cervical cytology (P.L., L.M., A.C., and C.F.) because they have been interpreting cervical cytology for the majority of their careers for >15 years. Conversely, 3 readers (C.D., F.M., and E.A.) had no experience in cervical cytology (2 medical students and 1 biologist with experience in molecular biology). All readers followed a 1-day course provided by the manufacturer, during which a training set of slides immunostained with p16/Ki-67 was presented. It was agreed that a positive result on the CINtec Plus kit would have been assigned if at least 1 cervical cell (either squamous or glandular) simultaneously demonstrated brown cytoplasmic and red nuclear staining (indicative of p16INK4a and Ki-67 expression, respectively). Cases without any evident double immunoreaction within the same cell (ie, demonstrating immunoreactivity for none or only 1 of the 2 markers) were termed negative. According to the 2001 Bethesda System, slides with <5000 well-preserved and well-visualized squamous epithelial cells were considered to be unsatisfactory for dual-staining evaluation.

After each reader had evaluated the first 150 slides, the cases that were found to have at least 2 different judgments were collegially discussed by the observers using a multihead microscope before reading the subsequent 350 slides. During such discussions, cytological and histological data were available for each woman and were considered. Original individual judgments remained unchanged for analysis. Dual-staining results were not used for the management of the patients.

Statistical Analysis

Slides judged as adequate as for dual staining by all readers were included in the analyses. Agreement between interpreters with regard to positivity (positive/negative) for dual staining was assessed by the kappa statistic for multiple raters.¹⁹ The standard error was obtained by bootstrapping.²⁰ Results were stratified by the interpreter's experience in cytology (expert vs nonexpert), the woman's baseline triage cytology, and whether the slide was interpreted before or after the consensus meeting.

The sensitivity and specificity of dual staining for CIN of type 2 or more severe histology (CIN2+) were estimated among HPV-positive women, considering all judgments. Women in the study who were positive for CIN2+ were those who had a histology of \geq CIN2 within 6 months of the same colposcopy in which the sample for dual staining was obtained (included). Women in the current study who were negative for CIN2+ were those with a histology <CIN2 (324 women) and those who did not have a biopsy taken (83 women). Because each slide was classified 7 times as positive or negative for dual staining, the sensitivity and specificity were estimated using methods for repeated measurements,²¹ in particular by logistic regression (in which each interpreter's judgment was the dependent variable and CIN2+ histology was the explanatory variable)²² using generalized estimating equations.²³ The estimated values represent the average sensitivity and specificity of the readers. A detailed supplementary statistical appendix is available online (see online supporting information).

Because slides from women with cytology that was within normal limits (WNL) were underrepresented compared with their frequency among all HPV-positive women, a kappa value for all HPV-positive women was estimated as a weighted mean of the kappa values observed among HPV-positive women with normal

cytology and among those with abnormal cytology (including unsatisfactory), with weights inversely proportional to sampling fractions (taking into account that, among women recruited in the Turin pilot project during the study period and who were HPV positive, 65% were judged to have negative cytology that was WNL). Specificity among all HPV-positive women was also estimated from the weighted mean of the relevant regression coefficients.

Analyses were performed using STATA statistical software (version 12; StataCorp, College Station, Tex) (kappa values) and SAS statistical software (version 9.2; SAS Institute Inc, Cary, NC) (generalized estimating equation models).

RESULTS

Of the 500 slides that were independently interpreted by 7 readers, 477 were determined to be satisfactory by all readers. Table 1 reports their distribution based on the cytology and histology. Distribution based on the number of judgments of positive for dual staining is reported in Table 1 in the Supporting Information (see online supporting information), and also was stratified by the presence of CIN2+.

Table 1. Study Slides by Woman's Triage Cytology and Worse Histology

Cytology	Total	Histology \geq CIN2
Within normal limits	114	3
Unsatisfactory	38	0
ASC-US	57	8
LSIL	210	18
ASC-H/HSIL1	58	41

Abbreviations: ASC-H, atypical squamous cells cannot exclude high-grade squamous intraepithelial lesion; ASC-US, atypical squamous cells of undetermined significance; CIN, cervical intraepithelial neoplasia; HSIL+, high-grade squamous intraepithelial lesion or more severe; LSIL, low-grade squamous intraepithelial lesion.

The kappa value was 0.66 (95% confidence interval [95% CI], 0.62-0.71) for women with triage cytology \geq ASC-US and 0.71 (95% CI, 0.63-0.79) for women with WNL cytology (Table 2). The difference was not statistically significant ($P = .30$). The estimated overall kappa value for all HPV-positive women was 0.70 (95% CI, 0.60-0.77). Kappa values for expert and nonexpert readers were found to be similar within each cytological category and there was no significant difference observed ($P = .88$ and $P = .11$, respectively, among women with normal cytology and \geq ASC-US).

Table 2. Kappa Values for Dual-Staining Positivity by Woman's Triage Cytology and Rater's Experience

	All Raters (95% CI)	Expert (95% CI)	Nonexpert (95% CI)	P for Expert Versus Nonexpert
ASC-US+ or unsatisfactory cytology	0.66 (0.62-0.71)	0.68 (0.62-0.74)	0.67 (0.59-0.76)	11
Within normal limits	0.71 (0.63-0.79)	0.74 (0.65-0.82)	0.73 (0.63-0.83)	88

Projected all HPV-positive women	0.70 (0.60-0.77)	0.73 (0.61-0.80)	0.72 (0.58-0.81)	-
----------------------------------	------------------	------------------	------------------	---

Abbreviations: 95% CI, 95% confidence interval; ASC-US+, atypical squamous cells of undetermined significance or more severe; HPV, human papillomavirus.

Conversely, agreement significantly increased ($P = .0066$) between the first set of slides, which were interpreted before the consensus discussion, and the second set, which were interpreted afterward (Table 3).

Table 3. Kappa Values for Dual-Staining Positivity by Period of Interpretation and Rater's Experience

	All Raters (95% CI)Expert	(95% CI)	Nonexpert (95% CI)
First group of slides (before discussion)	0.60 (0.51-0.67)	0.63 (0.55-0.72)	0.56 (0.46-0.67)
Second group of slides (after discussion)	0.72 (0.67-0.77)	0.73 (0.67-0.78)	0.75 (0.68-0.80)

Abbreviation: 95% CI, 95% confidence interval.

The specificity of dual staining for a histology <CIN2 was 65.0% among women with normal triage cytology and 59.1% among those with abnormal or unsatisfactory triage cytology (Table 4). The projected specificity of double staining among all HPV-positive women was 64.0% (95% CI, 57.4%-70.2%) for all readers, 66.7% (95% CI, 59.8%-73.0%) for expert readers, and 60.5% (95% CI, 59.8%-73.0%) for nonexpert readers. The sensitivity for a histology ≥CIN2 was 85.5% (95% CI, 77.9%-90.8%) among women with ASC-US or more severe/unsatisfactory triage cytology; in particular, the sensitivity was 85.8% (95% CI, 77.9%-91.2%) for expert readers and 85.1% (95% CI, 76.6%-90.9%) for nonexpert readers. Among women with WNL triage cytology, only 3 women were found to have a histology ≥CIN2 (all slides were judged as positive for dual staining by all raters). Thus, a projected sensitivity for all HPV-positive women was not estimated.

Table 4. Specificity and Sensitivity of Dual Staining for CIN2+ by Woman's Triage Cytology and Reader's Experience

	All Raters, % (95% CI)	Expert, % (95% CI)	Nonexpert, % (95% CI)	P for Expert Versus Nonexpert
Specificity for <CIN2, triage cytology ASC-US+, or unsatisfactory	59.1 (54.4-63.7)	61.1 (56.3-65.8)	56.4 (51.4-61.3)	.001
Specificity for <CIN2, triage cytology within normal limits	65.0 (61.6-68.4)	67.8 (63.4-72.1)	61.3 (56.0-66.5)	.059
Specificity for <CIN2 projected to all HPV-positive women	64.0 (57.4-70.2)	66.7 (59.8-73.0)	60.5 (59.8-73.0)	

Sensitivity for \geq CIN2, triage cytology ASC-US+, or unsatisfactory	85.5 (77.9-90.8)	85.8 (77.9-91.2)	85.1 (76.6-90.9)	.075
---	------------------	------------------	------------------	------

Abbreviations: 95% CI, 95% confidence interval; ASC-US+, atypical squamous cells of undetermined significance or more severe; CIN2+, cervical intraepithelial neoplasia of type 2 or more severe; HPV, human papillomavirus.

DISCUSSION

Dual p16INK4a/Ki-67 immunostaining was introduced mainly to obtain high reproducibility and specificity without the need to use morphologic criteria. Triage HPV-positive women is one of its most promising applications.¹¹ We are aware of only one other, very recent, study of the reproducibility of dual-staining interpretation.²⁴ When examining slides from HPV-positive women who underwent routine screening by HPV and cytology cotesting in the United States, the authors observed an overall kappa value of 0.71 between 11 expert cytotechnologists/pathologists, 10 of whom had only taken a short course on dual-staining interpretation. In the current study, based on slides from HPV-positive women screened by stand-alone HPV testing in Italy, we observed an overall kappa value of 0.70 between 7 readers, 3 of whom had no experience in cytology interpretation. Remarkably, especially after a short practice, they had reproducibility that was similar to that of experts in cervical cytology. To the best of our knowledge, comparable studies with stand-alone p16 staining are not available. Very high Kappa values of 0.84¹⁴ and 0.93¹² were reported previously, but in both cases only 2 expert raters, with strict previous collaboration, were considered.

We observed a specificity of 64%, which must be considered as good taking into account that dual staining was used to triage HPV-positive women. Much higher specificity is of course expected, and has been observed among the general screening population (95%).²⁵ With p16INK4a-only immunostaining, a specificity of 62% was observed among HPV-positive women.¹² However, in that study, as observed above, only 2 expert readers judged immunostaining. Indeed, the most relevant result from the current study is that, although specificity was slightly higher for interpreters who were experts in cytology compared with nonexperts, it remained high, even for the latter group (66.7% vs 60.5%). Similar specificity (63.9%) was observed among expert cytotechnicians/pathologists without experience in dual staining in the previously mentioned US study,²⁴ again with limited variability noted among observers.

We studied women invited to population-based screening who were reasonably representative of those women who will undergo HPV-based screening. Our estimates were for the triage of HPV-positive women. Observers were informed that women were HPV positive, as is expected to happen in practice. Underrepresentation of HPV-positive women with normal cytology was balanced by weighting. However, in this group, only those women who persisted in testing positive for HPV after 1 year underwent colposcopy. Compared with an HPV infection that was just acquired, such infections may have had a longer time to progress to overexpression of p16INK4a. This could have potentially resulted in underestimated specificity (sensitivity was not estimated for this group). To estimate reproducibility, we compared observers among them, thereby avoiding an external standard for dual-staining interpretation. Sensitivity and specificity estimate the average cross-sectional accuracy of the observers, thus taking into account the variability in the interpretation of dual staining. The data from the current study were obtained from samples stored in PreservCyt solution. However, the results are very similar to those obtained with SurePath (Becton, Dickinson and Company, Franklin Lakes, NJ), another commonly used liquid-based cytology transport medium.²⁴

The data from the current study suggest that, after a short training phase, the interpretation of dual staining could be performed even by staff not trained in the morphological interpretation of cytology. This finding is particularly interesting in settings in which skilled cytologists are not available in a sufficient number to implement screening based on HPV testing with cytological triage. This, for example, could be the case in low-income or middle-income countries. Training cytologists would require many years and entail a very high cost whereas training staff for dual-staining interpretation would be much faster and simpler.

FUNDING SUPPORT

This study received external financial support from Roche Diagnostics, which gave to the University of Turin the kits for double staining. The sponsor had no role in the design, data collection, data analysis, data interpretation, or writing of the article, or in the decision to submit for publication.

CONFLICT OF INTEREST DISCLOSURES

Dr. Maletta was supported by a grant from Roche Diagnostics for work performed as part of the current study.

REFERENCES

- 1 Naucier P, Ryd W, Tornberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med*. 2007;357:1589–1597.
- 2 Kitchener HC, Almonte M, Thomson C, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol*. 2009;10:672–682.
- 3 Rijkaart DC, Berkhof J, Rozendaal L, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol*. 2012;13:78–88.
- 4 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;11:249–257.
- 5 Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. *N Engl J Med*. 2009;360:1385–1394.
- 6 Ronco G, Dillner J, Elfstrom M, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet*. 2014;383:524–532.
- 7 von Knebel Doeberitz M. New markers for cervical dysplasia to visualize the genomic chaos created by aberrant oncogenic papillomavirus infections. *Eur J Cancer*. 2002;38:2229–2242.
- 8 Wentzensen N, von Knebel Doeberitz M. Biomarkers in cervical cancer screening. *Dis Markers*. 2007;23:315–330.
- 9 Klaes R, Friedrich T, Spitkovsky D, et al. Overexpression of p16(INK4A) as a specific marker for dysplastic and neoplastic epithelial cells of the cervix uteri. *Int J Cancer*. 2001;92:276–284.

- 10 Sano T, Oyama T, Kashiwabara K, Fukuda T, Nakajima T. Expression status of p16 protein is associated with human papillomavirus oncogenic potential in cervical and genital lesions. *Am J Pathol.* 1998;153:1741–1748.
- 11 Bergeron C, Ronco G, Reuschenbach M, et al. The clinical impact of using p16INK4a immunochemistry in cervical histopathology and cytology: an update of recent developments [published online ahead of print April 17, 2014]. *Int J Cancer.* doi: 10.1002/ijc.28900.
- 12 Carozzi F, Confortini M, Dalla Palma P, et al; New Technologies for Cervical Cancer Screening (NTCC) Working Group. Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol.* 2008;9:937–945.
- 13 Carozzi F, Gillio-Tos A, Confortini M, et al; NTCC Working Group. Risk of high-grade cervical intraepithelial neoplasia during follow-up in HPV-positive women according to baseline p16-INK4A results: a prospective analysis of a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol.* 2013;14:168–176.
- 14 Wentzensen N, Bergeron C, Cas F, Vinokurova S, von Knebel Doeberitz M. Triage of women with ASCUS and LSIL cytology: use of qualitative assessment of p16INK4a positive cells to identify patients with high-grade cervical intraepithelial neoplasia. *Cancer.* 2007;111:58–66.
- 15 Wentzensen N, Bergeron C, Cas F, Eschenbach D, Vinokurova S, von Knebel Doeberitz M. Evaluation of a nuclear score for p16INK4a-stained cervical squamous cells in liquid-based cytology samples. *Cancer.* 2005;105:461–467.
- 16 Tsoumpou I, Arbyn M, Kyrgiou M, et al. p16(INK4a) immunostaining in cytological and histological specimens from the uterine cervix: a systematic review and meta-analysis. *Cancer Treat Rev.* 2009;35:210–220.
- 17 Heatley MK. Ki67 protein: the immaculate deception? *Histopathology.* 2002;40:483.
- 18 Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol.* 2000;182:311–322.
- 19 Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York: John Wiley & Sons; 1981.
- 20 Lee J, Fung KP. Confidence interval of the kappa coefficient by bootstrap resampling. *Psychiatry Res.* 1993;49:97–98.
- 21 Ronco G, Biggeri A. Estimating sensitivity and specificity when repeated tests are performed on the same subject. *J Epidemiol Biostat.* 1999;4:329–336.
- 22 Coughlin SS, Trock B, Criqui MH, Pickle LW, Browner D, Tefft MC. The logistic modelling of sensitivity, specificity, and predictive value of a diagnostic test. *J Clin Epidemiol.* 1992;45:1–7.
- 23 Zeger SL, Liang KY. An overview of methods for the analysis of longitudinal data. *Stat Med.* 1992;11:1825–1839.

- 24 Wentzensen N, Fetterman B, Tokugawa D, et al. Interobserver reproducibility and accuracy of p16/Ki-67 dual-stain cytology in cervical cancer screening [published online ahead of print August 12, 2014]. *Cancer (Cancer Cytopathol)*. doi: 10.1002/cncy.21473.
- 25 Ikenberg H, Bergeron C, Schmidt D, et al; PALMS Study Group. Screening for cervical cancer precursors with p16/Ki-67 dual-stained cytology: results of the PALMS study. *J Natl Cancer Inst*. 2013;105:1550–1557.