Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: a nested substudy of the NTCC randomised controlled trial

*Francesca Carozzi, Massimo Confortini, Paolo Dalla Palma, Annarosa Del Mistro, Anna Gillio-Tos, Laura De Marco, Paolo Giorgi-Rossi, Giovanni Pontenani, Stefano Rosso, Cristina Sani, Catia Sintoni, Nereo Segnan, Manuel Zorzi, Jack Cuzick, Raffaella Rizzolo, Guglielmo Ronco, and the New Technologies for Cervival Cancer screening (NTCC) working group**

Summary

Background **Human-papillomavirus (HPV) testing is more sensitive, but less specific, than conventional cytology for detecting high-grade cervical intraepithelial neoplasia (CIN). We assessed whether HPV testing with triage by p16-INK4A overexpression can increase specificity while maintaining sensitivity.**

Methods **HPV-positive women were enrolled between June 10, 2003, and Dec 31, 2004 in a multicentre randomised controlled trial, which compared stand-alone HPV testing by Hybrid Capture 2 (experimental group) with conventional cytology, were referred for colposcopy. In seven of nine centres, cytospin preparations from these women were tested for p16-INK4A overexpression by immunostaining. The sensitivity and specificity for CIN grade 2 or more, determined at blind review of histology, were calculated for these women. We also estimated the relative sensitivity and relative referral to colposcopy that would have been obtained by HPV testing with p16-INK4A triage compared with conventional cytology. This trial is registered as a Standard Randomised Controlled Trial, number ISRCTN81678807.**

Findings **24 661 women were randomly assigned to the experimental group. 1137 women (74% of those undergoing colposcopy in relevant centres), including 50 with CIN2 and 42 with CIN3 or cancer, had valid p16-INK4A immunostaining. For the endpoint of CIN2+, sensitivity and specificity of p16-INK4A (deemed positive with any number of stained cells—except endocervical, metaplastic, and atrophic cells if morphologically normal) in HPV-positive women of any age were 88% (81 of 92; 95% CI 80–94) and 61% (633 of 1045; 57–64), respectively. In the 35–60-year age group, the relative sensitivity of HPV testing and p16-INK4A triage versus conventional cytology for CIN2+ was 1·53 (95% CI 1·15–2·02) and relative referral was 1·08 (0·96–1·21). In the 25–34-year age group, relative sensitivity was 3·01 (1·82–5·17) and relative referral was 1·15 (0·96–1·37). In the latter age group, when 5% or more stained cells were deemed positive, the corresponding values were 2·06 (1·20–3·68) and 0·58 (0·46–0·73), respectively.**

Interpretation **HPV testing with p16-INK4A triage produces a significant increase in sensitivity compared with conventional cytology, with no substantial increase in referral to colposcopy.**

Funding **European Union, Italian Ministry of Health, Regional Health Administrations of Piemonte, Tuscany, Veneto, and Emilia-Romagna, and Public Health Agency of Lazio Region.**

Published **Online** September 9, 2008 DOI:10.1016/S1470-2045(08)70208-0

*Members listed at end

Centre for Cancer Study and Prevention, Florence, Italy (F Carozzi PhD, M Confortini PhD, G Pontenani BS, C Sani MSc); **S Chiara Hospital, Trento, Italy** (P Dalla Palma MD); **Venetian Oncology Institute—IRCCS, Padova, Italy** (A Del Mistro MD, M Zorzi MSc); **Centre for Cancer Prevention, Piemonte, Turin, Italy** (A Gillio-Tos PhD, L De Marco PhD, S Rosso MD, N Segnan MD, R Rizzolo MSc, G Ronco MD); **Laziosanità—Agency for Public Health, Lazio Region, Rome, Italy** (P Giorgi-Rossi PhD); **Ravenna Hospital, AUSL Ravenna, Italy** (C Sintoni PhD); **and Queen Mary's School of Medicine and Dentistry, and Cancer Research UK, London, UK** (Prof J Cuzick PhD)

Correspondence to: Dr Guglielmo Ronco, Unit of Cancer Epidemiology, Centre for Cancer Prevention, Piemonte, 10123 Turin, Italy **guglielmo.ronco@cpo.it**

Introduction

Several studies have shown that human papillomavirus (HPV) DNA testing as a primary screening test has greater sensitivity than conventional cytology for identifying high-grade cervical intraepithelial neoplasia (CIN).[1–4] For example, in a recent pooled analysis of European and north American studies (age range 18–69 years), HPV testing showed a pooled sensitivity for CIN grade 2 or above of 96% (95% CI 94–97) versus 53% (49–57) for cytology.[4] In the New Technologies for Cervical Cancer screening (NTCC) study,[1] the relative sensitivity of HPV testing by Hybrid Capture 2 (HC2) at 1 pg/mL DNA cut-off versus conventional cytology was 1·63 (95% CI 1·25–2·12) in women aged 35–60 years. However, HPV testing is less specific than cytology because many infections regress without developing high-grade lesions. Loss in specificity is restricted in absolute terms. For example, in the above mentioned pooled analysis, HPV testing showed a pooled specificity of 91% (95% CI 90–91) versus 96% (96–97) for cytology.[4] Nevertheless, this loss in specificity has a relevant effect in terms of positive predictive value (PPV), in view of the low prevalence of high-grade lesions in well-screened women in many populations. In the NTCC study, the relative PPV of HC2 versus conventional cytology was 0·67 (95% CI 0·52–0·87) in women aged 35–60 years at the 1 pg/mL cut-off.[1] There is therefore a need to identify strategies for increasing specificity with HPV DNA testing while maintaining its advantage in terms of sensitivity.

p16-INK4A, hereafter denoted as p16, is a cyclin-dependent kinase inhibitor, the expression of which is negatively controlled by the *RB1* gene product. Usually, p16 is expressed at very low concentrations in healthy cells,

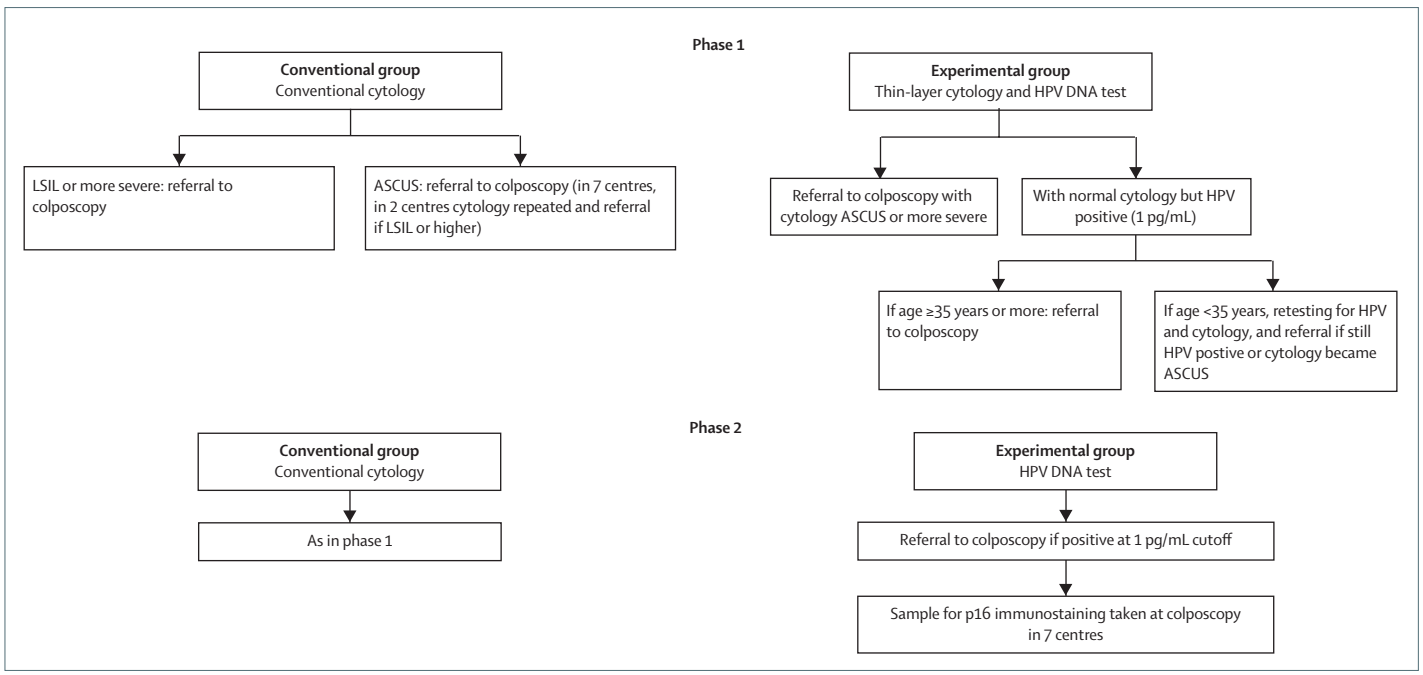

Figure 1: **NTCC study protocol for testing and management according to group and phase**
LSIL=low-grade squamous intraepithelial lesion. ASCUS=atypical cells of undetermined significance. HPV=human papillomavirus. p16=p16-INK4A.

whereas it is strongly over expressed in cervical-cancer cell lines in which *RB* has been functionally inactivated by the high-risk HPV E7 oncoprotein.[5,6] Therefore, p16 overexpression, identified by immunostaining or enzyme-linked immunosorbent assay (ELISA), can be considered as a marker, not only of HPV infection, but also of activated expression of viral oncogenes and of virus-induced deregulation of the cell cycle,[7,8] making it a candidate test for triaging HPV-DNA-positive women.

We did a study, nested in the NTCC study, to estimate the sensitivity and specificity of p16 overexpression in HPV-DNA-positive women for histologically confirmed CIN grade 2 or 3 or cervical cancer (CIN2+) and for CIN grade 3 or cervical cancer (CIN3+).

Methods

Patients and procedures

Methods applied during the NTCC study are described in detail elsewhere.[1,9,10] Briefly, after providing written informed consent, women aged 25–60 years attending for a new round of routine cervical-cancer screening in nine Italian organised screening programmes, were randomly assigned to two study groups, conventional (conventional cytology) and experimental (HPV testing alone or with liquid-based cytology), during two subsequent phases (figure 1). Women were excluded if they were pregnant, a virgin, had undergone hysterectomy, or had been treated for CIN or cancer during the past 5 years. About half of the enrolled women (23 680 of 47 127 [50%] in the conventional group and 24 021 of 47 243 [51%] in the experimental group) had undergone a screening test registered in organised programmes during the previous 4 years. The trial was approved by the local ethics committees of participating centres.

During phase 2,[1] between June 10, 2003, and Dec 31, 2004, women who were randomly assigned to the experimental group underwent HPV DNA testing alone, which was done in seven laboratories by use of Hybrid Capture 2 (HC2; Digene Corporation, Gaithersburg, MD, USA). Only probes designed to detect high-risk HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68 were used. HC2 results were expressed as the ratio of the specimen's light emission compared with the mean of three concurrently tested controls of 1 pg/mL HPV DNA (ie, 1 relative light unit [RLU]=1 pg/mL of HPV DNA in the specimen). A high reproducibility between participating laboratories (Kappa=0·93 for positive *vs* negative with standard transport medium) was noted when circulating clinical samples.[11] Women were directly referred to colposcopy if the HPV test was positive with a 1·0 RLU cut-off. In a previous phase (phase 1), women were tested for both HPV DNA and liquid-based cytology, and those who were HPV positive were directly referred to colposcopy if aged 35–60 years,[9] or were triaged by cytology if aged 25–34 years.[10] Women who were randomly assigned to the conventional group had conventional cytology that was interpreted by cytoscreeners in 14 local laboratories participating in regular screening programmes and classified according to the Bethesda 1991 system.[12] Women in the conventional group were managed according to the regular protocol of each centre.[1] The protocol by phase and group is summarised in figure 1.

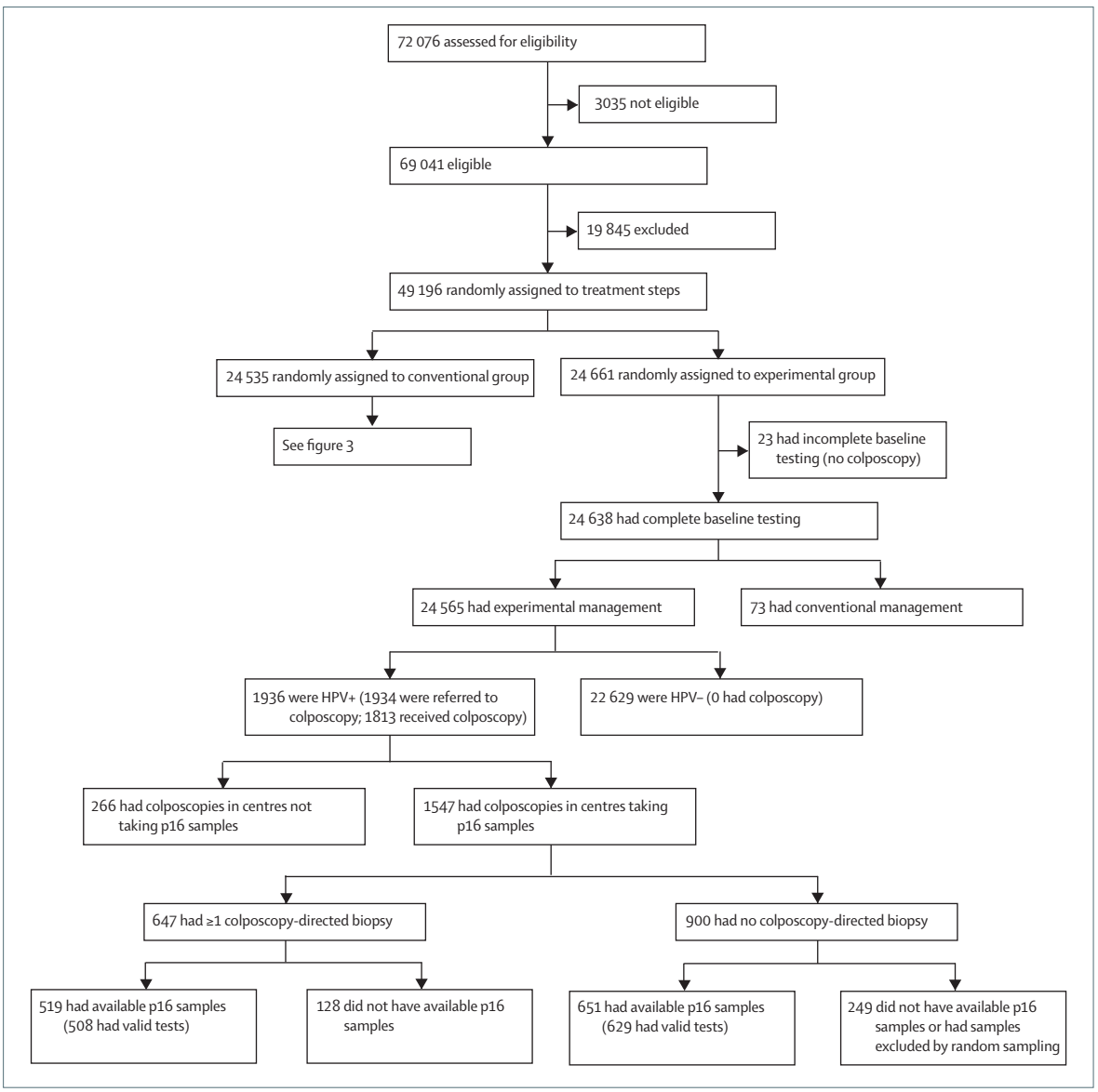

Figure 2: **Nested substudy profile for experimental group**
HPV=human papillomavirsus. p16=p16-INK4A.

The same colposcopists examined women in the experimental group and in the conventional group. They had access to patient's notes, both for cytology and HPV. Suspicious areas were biopsied.

Specimens for p16 staining were collected from HPV-positive women at the time of the first colposcopy in seven of the nine study centres. The two remaining centres did not deem this feasible. There was no heterogeneity between these two centres and the others in terms of relative sensitivity (p=0·34) and relative proportion of referrals to colposcopy (p=0·59) of HC2 versus conventional cytology. For organisational reasons, specimen collection started at different times in each centre, with some delay from the start of phase 2. To contain costs, we aimed at testing all women with CIN2+ and only a sample of 1100 women without CIN2+. This size was computed so as to have confidence intervals for specificity of ±3% in the least precise situation of 50% true specificity. Because the wait for the histology result could have compromised the preservation of material, we sampled specimens from women who had no biopsy. In practice, in five of seven centres a systematic random sample (by colposcopy order) of 20% of specimens from women who had had no biopsy taken was discarded.

The main endpoints were sensitivity and specificity for CIN2+ and the relative sensitivity and relative referal to colposcopy obtained with HPV testing and p16 triage, compared with conventional cytology.

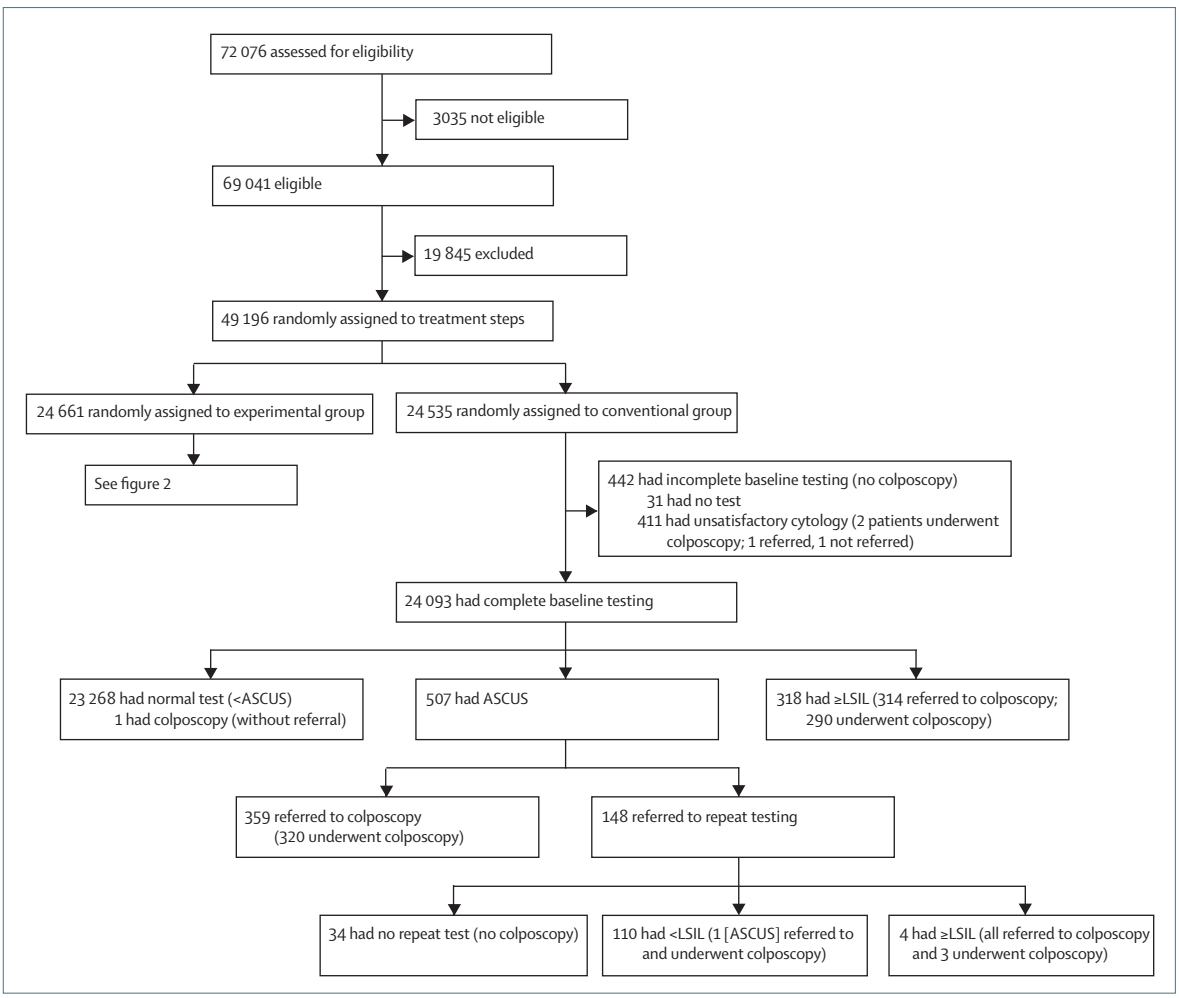

Figure 3: **Nested substudy profile for conventional group**
ASCUS=atypical cells of undetermined significance. LSIL=low-grade squamous intraepithelial lesion.

p16 immunostaining

Cervical material was collected, before colposcopy, in a ThinPrep vial containing PreservCyt transport medium (Cytyc Corporation, Boxborough, MA, USA) by use of a plastic Ayre's spatula and cytobrush. For p16 immunostaining, after preparation of one slide for cytology, 2 mL of the residual Thin-Prep fluid for each specimen was used for a cytospin preparation. The choice of cytospin preparation was aimed at a more efficient use of the residual ThinPrep fluid. In a previous study,[13] similar findings were obtained on cytospin preparations and thin-layer slides. Positive and negative control slides, prepared from a cell pool containing residual clinical samples with high-grade squamous intraepithelial lesions (HSIL) and negative cytological findings, respectively, were treated in the same way and used in each immunostaining procedure. After cytocentrifugation (5 min at 209 G), slides were air-dried for 10 min, then treated with spray-fixation reagent, containing polyethylene glycol, and immunostained within 24 h. Before beginning the immunostaining procedure, all spray-fixed specimens were incubated in 50% volume/volume (v/v) alcohol, followed by one washing step in deionised water. For immunostaining, we used a p16–INK4A-specific monoclonal antibody (clone E6H4) and the CINtec p16-INK4A Cytology kit (Dako Cytomation, now Dako A S, Glostrup, Denmark) was applied, as reported previously.[13] In brief, smears were treated with 3% (v/v) hydrogen peroxide and then submitted for epitope retrieval at 95–99°C for 40 min (±1 min). After cooling, the p16 antibodies were applied for 30 min (±1 min) and then a reagent for observation and substrate-chromogen solutions were added. Haematoxylin was used as a counterstain.

Cytospin preparation and staining were centralised. To assess the cellularity of each slide, all cells in five different fields (at random) were counted (objective ×40) and the total number of cells on the slide was assessed as the mean number of cells per microscopic field multiplied by the total number of fields. For the purpose of this

study, we deemed slides with less than 500 cells as inadequate. A p16 negative finding was defined as no cell staining or staining of just morphologically normal endocervical, metaplastic, or atrophic cells. Furthermore, staining of bacteria was not deemed a positive staining. All the remaining immunostained cells, including superficial, intermediate, and parabasal normal cells and all abnormal cells in the full cytospin surface area, were counted and the fraction of positive cells was calculated and further classified according to the proportion of stained cells: less than 5%, 5% to less than 10%, and 10% or more, respectively. Slides were independently read by two investigators blind to cytological and histological diagnoses. The few (41 of 1170 [4%]) discordant readings were resolved by consensus review.

This NTCC trial is registered as an International Standard Randomised Controlled Trial, number ISRCTN81678807.

Statistical analyses

We estimated the sensitivity and specificity of p16 overexpression in HPV-positive women for histologically confirmed CIN2+ and CIN3+. We considered all histological specimens (including surgical specimens from an ablative treatment) collected within 1 year from referral to colposcopy. These specimens were blindly and independently reviewed as previously described.[14] The most severe diagnosis for each woman was used. Because different proportions of women with and without biopsy were tested, specificity was computed as a weighted mean of the proportion of p16-negative women who had no biopsy and who had biopsy <CIN2 (or <CIN3, according to the endpoint considered), with weights inversely proportional to the number tested (considering all centres). Confidence intervals were computed on the basis of the exact binomial distribution for sensitivity and of its normal approximation for specificity.

We also estimated the relative sensitivity and the relative referral—compared with conventional cytology—that would have been obtained by only referring to colposcopy the HPV-positive women who were also positive to p16 immunostaining. Relative sensitivity was computed as the product of the relative sensitivity for HPV testing alone—previously estimated[1] as the ratio of detection rates of CIN2+ (or CIN3+) in the two NTCC groups[15]—and the sensitivity of p16 in HPV-positive women obtained in the present study. Relative referral was computed as the product of the relative referral that would have been obtained with direct referral to colposcopy of all HPV-positive women and the estimated proportion of all HPV-positive women who were also p16 positive. The latter was calculated as a weighted mean of the proportion of women who were p16 positive in each category of histology (ie, no biopsy, biopsy <CIN2, CIN2+). We analysed women in two age groups (25–34 years and 35–60 years). For the older age group, we used the pooled estimates of relative sensitivity and referral for HPV testing obtained by combining the findings of both phases of recruitment.[1] Indeed, no evidence of heterogeneity of relative sensitivity between phases or centres was noted for this age group.[1] For the younger age group, we used only the data from phase 2, because for phase 1 we had only estimates for HPV triaged by cytology.[10]

	1+ cells stained, n (%)	≥5% cells stained, n (%)	≥10% cells stained, n (%)	Total number of women who had a p16 test
All ages				
No biopsy	226 (36)	130 (21)	54 (9)	629
Biopsy (no CIN)	94 (39)	48 (20)	21 (9)	241
CIN1	92 (53)	33 (19)	11 (6)	175
CIN2	43 (86)	29 (58)	18 (36)	50
CIN3+	38 (91)	27 (64)	17 (41)	42
Age 25–34 years				
No biopsy	81 (29)	42 (15)	12 (4)	277
Biopsy (no CIN)	47 (33)	22 (15)	9 (6)	143
CIN1	50 (51)	26 (26)	10 (10)	99
CIN2	26 (79)	17 (52)	11 (33)	33
CIN3+	19 (95)	14 (70)	8 (40)	20
Age 35–60 years				
No biopsy	145 (41)	88 (25)	42 (12)	352
Biopsy (no CIN)	47 (48)	26 (27)	12 (12)	98
CIN1	42 (55)	7 (9)	1 (1)	76
CIN2	17 (100)	12 (71)	7 (41)	17
CIN3+	19 (86)	13 (59)	9 (41)	22

Table 1: **Proportion of women with positive p16-INK4A (p16) immunostaining at different cut-offs, by histology**

Because the variance of the compound indexes of relative sensitivity and relative referral was not easily analytically derivable to our knowledge, we resorted to the Monte Carlo Markov Chain (MCMC) method to estimate these parameters and their distribution. By considering the present part of the study as being done on a subsample of the larger study population, we used the empirical previous knowledge from the whole study findings. We therefore calculated the 95% CIs of both relative sensitivity and relative referral from their "a-posteriori" distributions, after sampling from beta distributions with parameters from the observed values (eg, positive cases over total tested cases of the pertinent group) of each proportion. The median of the distribution was used as point estimate. Sampling was done with two MCMC runs with 10 000 cycles by use of the WinBUGS 1.4.3 software.[16]

Role of the funding source

This study was funded by the European Union, the Italian Ministry of Health, Regional Health Administrations of Piemonte, Tuscany, Veneto, and Emilia-Romagna, and the Public Health Agency of Lazio Region. The funding sources had no role in the study

	Endpoint CIN2+		Endpoint CIN3+	
	Sensitivity (95% CI), n %	Specificity* (95% CI), n %	Sensitivity (95% CI), n %	Specificity* (95% CI), n %
All ages				
1+ cells stained	81/92 (88; 80–94)	633/1045 (61; 57–64)	38/42 (91; 77–97)	640/1095 (59; 55–63)
≥5% cells stained	56/92 (61; 50–71)	834/1045 (79; 76–83)	27/42 (64; 48–78)	855/1095 (78; 75–81)
≥10% cells stained	35/92 (38; 28–49)	959/1045 (91; 89–94)	17/42 (41; 26–57)	991/1095 (90; 88–93)
Age 25–34 years				
1+ cells stained	45/53 (85; 72–93)	341/519 (66; 60–72)	19/20 (95; 75–99·9)	348/552 (63; 57–69)
≥5% cells stained	31/53 (59; 44–72)	429/519 (83; 78–87)	14/20 (70; 46–88)	445/552 (81; 76–85)
≥10% cells stained	19/53 (36; 23–50)	488/519 (94; 91–97)	8/20 (40; 19–64)	510/552 (92; 89–96)
Age 35–60 years				
1+ cells stained	36/39 (92; 79–98)	292/526 (57; 51–63)	19/22 (86; 65–97)	292/543 (56; 50–61)
≥5% cells stained	25/39 (64; 47–79)	405/526 (76; 71–81)	13/22 (59; 36–79)	410/543 (75; 70–80)
≥10% cells stained	16/39 (41; 26–58)	471/526 (89; 85–93)	9/22 (41; 21–64)	481/543 (88; 85–92)

*Specificity was computed as a weighted mean of the proportion of p16-negative women in those without biopsy and in those with biopsy findings of <CIN2 when CIN2+ is the endpoint and of <CIN3 when CIN3+ is the endpoint, with weights inversely proportional to the fraction tested.

Table 2: **Sensitivity and specificity of p16-INK4A (p16) immunostaining in HPV-DNA-positive women by age and cut-off**

	Relative sensitivity for CIN2+ (95% CI)	Relative sensitivity for CIN3+ (95% CI)	Relative referral to colposcopy (95% CI)
Age 25–34 years			
HPV testing (≥1 pg/mL) with no triage	3·50 (3·11–5·82)*	2·61 (1·21–5·61)*	3·64 (3·17–4·19)
HPV testing (≥1 pg/mL) and p16 triage (1+ cells stained)	3·01 (1·82–5·17)	2·52 (1·18–5·78)	1·15 (0·96–1·37)
HPV testing (≥1 pg/mL) and p16 triage (≥5% cells stained)	2·06 (1·20–3·68)	1·84 (0·81–4·38)	0·58 (0·46–0·73)
Age 35–60 years			
HPV testing (≥1 pg/mL) with no triage	1·63 (1·25–2·12)*	1·52 (1·06–2·19)*	2·38 (2·21–2·57)
HPV testing (≥1 pg/mL) and p16 triage (1+ cells stained)	1·53 (1·15–2·02)	1·32 (0·88–1·95)	1·08 (0·96–1·21)
HPV testing (≥1 pg/mL) and p16 triage (≥5% cells stained)	1·06 (0·73–1·49)	0·90 (0·52–1·45)	0·55 (0·46–0·64)

*Data from reference 1. p16=p16-INK4A.

Table 3: **Relative sensitivity and relative referral for different HPV-based screening strategies versus conventional cytology**

design, data collection, data analysis, data interpretation, or writing the report. The members of the NTCC steering committee (FC, MC, PDP, PGR, NS, MZ, JC, and GR) had full access to all the data in the study. GR had final responsibility for the decision to submit for publication.

Results

During the second phase of recruitment of the NTCC trial, 24 661 eligible women (median age 42 years [range 25–60 years]) were randomly assigned to the experimental group. 1936 of these women were HPV positive, 1813 (94%) of whom had a colposcopy.

A p16 immunostaining test was done for 1170 of 1547 women (76%) who underwent colposcopy in the seven centres that collected specimens. In 33 of 1170 women (3%; 95% CI 2–5), the test was deemed unsatisfactory, because the samples were obscured by blood, inflammatory exudates, or had insufficient cellularity. Thus, 1137 women had a valid p16 test (508 of 647 [79%] women with biopsy and 629 of 900 [70%] women without biopsy). The distribution of histology was very similar between women with valid or unsatisfactory p16 tests (data not shown). Of the 377 women who did not have a p16 test, 124 women without biopsy had been sampled for exclusion in the five centres that did this. In four patients, the aliquot for p16 was not taken because less than 2 mL of ThinPrep fluid were available, whereas in 249 women no specimen was available because of problems of organisation. The study flowchart is shown in figures 2 and 3.

The unweighted kappa value between the two interpreters on the four categories (no staining, <5%, 5% to <10%, ≥10% cell stained) was 0·93 (95% CI 0·91–0·94).

When considering all ages, the proportion of patients who showed p16 overexpression was similar in HPV-positive women who had no biopsy (226 of 629 [36%]) and in those who had biopsy but no evidence of CIN (94 of 241 [39%]). In women with CIN, this proportion increased with CIN grade, from 53% (92 of 175) for CIN1 to 91% (38 of 42) for CIN3 or invasive cancer. The proportion with 5% or more stained cells was similar in women with no biopsy, biopsy but no CIN, and CIN1 (table 1).

Sensitivity and specificity for CIN2+ and CIN3+ of p16 immunostaining in HPV-positive women, with different cut-off values, are reported in table 2. When considering all ages, sensitivity was 88% (81/92; 95% CI 80–94) and specificity 61% (633/1045; 57–64) with CIN2+ as the endpoint and when the test was defined as positive in the presence of any stained cell. With CIN3+ as the endpoint, the corresponding values were 91% (38/42; 77–97) and 59% (640/1095; 55–63), respectively. When increasing the cut-off to 10% or more stained cells with CIN2+ as the endpoint, specificity increased to 91% (959/1045; 89–94), but sensitivity decreased to 38% (35/92; 28–49).

In the 35–60-year age group, the relative sensitivity for CIN2+ of HPV without p16 triage (ie, when referring to colposcopy in all HPV-positive women) versus conventional cytology was 1·63 (1·25–2·12),[1] but referral was more than doubled (relative referral 2·38; 2·21–2·57; table 3). If HPV-positive women were triaged by p16, considering any stained cells as positive, the relative sensitivity only slightly decreased to 1·53 (1·15–2·02), but referral was similar to that for conventional cytology (relative referral 1·08; 0·96–1·21). When positivity was defined as 5% or more stained cells, referral further decreased (relative referral 0·55; 0·46–0·64), but there was no gain in terms of sensitivity compared with cytology (relative sensitivity 1·06; 0·73–1·49; table 3).

For the 25–34-year age group, the relative sensitivity versus cytology was 3·50 (3·11–5·82) without triage,[1] 3·01 (1·82–5·17) with p16 triage and any positive cell defined as the cut-off, and 2·06 (1·20–3·68) with 5% or more stained cells as the p16 cut-off. The corresponding relative referral was 3·64 (3·17–4·19), 1·15 (0·96–1·37), and 0·58 (0·46–0·73), respectively.

Discussion

In our study, HPV-positive women who were also positive for p16 overexpression, maintained almost all the gain in sensitivity obtained by HPV testing without p16 triage, by comparison with conventional cytology, but with referral to colposcopy similar to that of conventional cytology.

p16 immunostaining has previously been assessed as a method for improving the interpretation of histology[17–22] or for triaging women with cytology of ASCUS or LSIL.[13,23–29] In a study of women referred to colposcopy because of abnormal cytology, p16 ELISA showed similar sensitivity and slightly better specificity than HPV testing by Hybrid Capture 2.[30] To our knowledge, our study is the first assessment of p16 testing for triaging unselected HPV-positive women.

A note should be made that the sensitivity and specificity estimates reported here are calculated in HPV-positive women and apply to these patients only, and not to the general population in which specificity is plausibly much higher. Furthermore, we used cytospin preparations for immunostaining and, therefore, our findings apply to this preparation. More cells are present on standard liquid-based cytology slides, which could potentially increase sensitivity and decrease specificity. However, the 88% sensitivity for CIN2+ that we noted in HPV-positive women was comparable to that recorded in women with abnormal cytology, ranging from 78%[27] to 96%.[28]

The present data do not allow direct estimation of the persistence of lesions detected by HPV testing with p16 triage. p16 overexpression has been shown to be associated with progression to CIN3 or cancer.[31,32] Our previous data[1] suggest that at a younger age, direct referral to colposcopy for all HPV-positive women could lead to the detection of regressive lesions. During phase 1 of the NTCC trial, in patients aged 25–34 years who had cytology triage and 1-year follow-up, relative sensitivity for CIN2+ versus cytology alone was 1·6.[10] In this substudy nested in phase 2, relative sensitivity with p16 triage in women of the same age versus cytology alone was 3·01 with any number of stained cells and 2·06 with 5% or more stained cells (table 3). These findings suggest that in younger women, parts of p16-positive lesions could also be regressive if the percentage of cells overexpressing p16 is low. Despite its potential usefulness in this age group, consideration of the proportion of stained cells was shown to be of little practical use.

A direct comparison between cytology and p16 as methods for triaging HPV-positive women is not currently possible. During the second phase of the NTCC trial, women in the experimental group were screened by HPV-testing only. Liquid-based cytology was done at colposcopy, but its interpretation was not blinded to histological findings.

Increased specificity with p16 has also been obtained by use of the nuclear features of stained cells.[26] This method is yet to be used in our study. However, we did a morphological assessment, because endocervical, metaplastic, and atrophic normal cells, which can have p16 immunoreactivity, were considered negative, even if stained, which plausibly increased specificity. The use of this nuclear score is crucial if p16 is applied as a triage method for abnormal cytology or as a primary screening test, but it might be less relevant in women already selected for being HPV positive. Indeed, our data show that in HPV-positive women p16 overexpression is strongly associated with the presence of histologically confirmed CIN2+, suggesting that it actually is a marker of progression.

In this nested substudy, all HPV-positive women were referred for colposcopy and compliance with colposcopy was high. Furthermore, we obtained specimens for p16 staining in a large proportion of eligible women. Additionally, cases without p16 findings were plausibly at random. The use of different sampling fractions in women with and without biopsy was balanced by weighting for the reciprocal of the fraction tested. The primary end point, CIN2+, was determined by a histological review blinded to HPV test, cytology, and p16 findings. However, some misclassification is still

possible, especially for CIN2, which might have led to a slight underestimation of the accuracy of p16. Random biopsies in addition to colposcopy-directed biopsies have been shown to increase the sensitivity for high-grade CIN.[33] We considered all histology obtained within 1 year from referral, including surgical specimens, to decrease the proportion of high-grade lesions potentially missed. One study showed that high-grade CIN detected by random colposcopies did not affect the sensitivity estimates for cytology and HPV testing.[34] However, an effect on p16 accuracy cannot be completely excluded. Colposcopists were not blinded to the findings of cytology and HPV testing (but were blinded to the finding of p16 assessment, done after colposcopy). The mean number of colposcopies (per woman who had at least one, mean 1·34 [SD 0·57] in the conventional group *vs* 1·21 [0·65] in the experimental group) and the proportion of biopsies (323 of 617 [52%] of women who underwent colposcopy had biopsy in the conventional group *vs* 788 of 1813 [43%] in the experimental group) were slightly higher in the conventional group.[1] Therefore, this could, at most, have led to a slightly underestimated relative sensitivity of HPV with p16 triage versus cytology.

The study was population-based and part of a routine organised screening activity in a low-risk population. More than 70% of eligible women were enrolled,[1] suggesting that the findings are applicable to routine practice.

This study supports the application of triage by p16 immunostaining in HPV-positive women. Refining the criteria for positivity, either by use of higher cut-offs in younger women or nuclear scoring, might further improve performance.

In the NTCC study, HPV-positive women, independent of p16 result, are retested by HPV and cytology at yearly intervals. This testing will allow us to study whether the presence of p16 overexpression in HPV-positive women predicts the future development of high-grade intraepithelial lesions. The risk of these lesions in HPV-positive, but p16-negative, women will determine if they need short-interval retesting or if they can return to routine screening.

Our findings were obtained with women who had not received prophylactic HPV vaccination and, therefore, are not immediately applicable to vaccinated women. Genotyping of the HPV infections detected is currently being done. p16 accuracy in women with HPV infection from non-16/18 types will provide a proxy for vaccinated women.

Contributors

GR was the NTCC project leader and designed this study with FC, AG-T, and ADM. GR, FC, and JC drafted the manuscript. CSa and GP did the immunostaining procedure under the supervision of FC. MC and FC assessed the immunostained slides. AG-T, LDM, ADM, and CSi provided the cervical-cell samples. GR, SR, and RR did the data analysis. MC, PDP, PG-R, NS, and MZ organised the local fieldwork. All authors critically revised the manuscript.

NTCC Working Group Collaborators in Italy

Turin—P Giubilato, B Ghiringhello, F Parisio, E Berardengo, D Campione, S Coverlizza, S Taraglio, M G Accinellli, R Volante; *Trento*—E Polla, A Pojer, S Girlando, D Aldovini; *Veneto*—M Vettorazzi, D Minucci, M Matteucci, L Onnis, E Insacco, M Lestani, A Vignato; *Emilia-Romagna*—C Naldoni, M Manfredi, P Pierotti, G P Casadei, G Collina, M Serafini, B Vitali, M Aldi, S Folicaldi, R Nannini, G Galanti, M De Lillo; *Florence*—A Iossa, S Ciatto, M P Cariaggi, S Cecchini, M Zappa, G L Taddei; *Lazio*—S Brezzi, P Raggi, E Gomes, A Pellegrini, M L Schiboni, P Borgia, F Chini.

Conflicts of interest

JC's institution has received grants from MTM Laboratories (Heidelberg, Germany) and Quiagen (Hilden, Germany) for other studies. Minor payment for participation in advisory boards for firms whose products are not involved in the present study was received by FC (Abbot, Roche, and GenProbe) and GR (GenProbe). All other authors declared no conflicts of interest.

Acknowledgments

This study was financially supported by the European Union (Europe against Cancer contracts SI.2.327046 and SPC.2002475), by the Italian Ministry of Health (Special Projects "Valutazione di nuove tecnologie per lo screening del cervicocarcinoma" and "Valutazione di nuove tecnologie per lo screening del cervicocarcinoma—follow-up" and L 138/2004 project "NTCC—Nuove tecnologie per lo screening del cervicocarcinoma"), by Regional Health Administrations of Piemonte (Project "Nuove Tecnologie per lo screening del cervicocarcinoma"), Tuscany, Veneto, and Emilia-Romagna, and by the Public Health Agency of Lazio Region (Project "Valutazione di nuove tecnologie per lo screening del cervicocarcinoma"). We thank all the staff for their assistance in running the study and the thousands of women who have participated in this study.

References

1 Ronco G, Giorgi-Rossi P, Carozzi F, et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone to conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst* 2008; **100:** 492–501.

2 Mayrand MH, Duarte-Franco E, Rodrigues I, et al. Human papillomavirus DNA versus Papanicolau screening tests for cervical cancer. *N Engl J Med* 2007; **357:** 1579–88.

3 Cuzick J, Szarewski A, Cubie H, et al. Management of women who test positive for high-risk types of human papillomavirus: the HART study. *Lancet* 2003; **362**: 1871–76.

4 Cuzick J, Clavel C, Petry C-H, et al. Overview of the European and north American studies on HPV testing in primary cervical screening. *Int J Cancer* 2006; **119:** 1095–101.

5 Khleif SN, DeGregori J, Yee CL, et al. Inibition of cyclin D-CDK4/CDK6 activity is associated with E2F-mediated induction of cyclin kinase inibitor activity. *Proc Natl Acad Sci USA* 1996; **93:** 4350–54.

6 Nakao Y, Yang X, Yokoyama M, et al. Induction of p16 during immortalization by HPV 16 and 18 and not during malignant transformation. *Br J Cancer* 1997; **75:** 1410–16.

7 von Knebel Doeberitz M. New molecular tools for efficient screening of cervical cancer. *Dis Markers* 2001; **17:** 123–28.

8 von Knebel Doeberitz M. New markers for cervical dysplasia to visualise genomic chaos created by aberrant oncogenic papillomavirus infection. *Eur J Cancer* 2002; **38:** 2229–42.

9 Ronco G, Segnan N, Giorgi-Rossi P, et al. Human papillomavirus testing and liquid-based cytology: results at recruitment from the New Technologies for Cervical Cancer randomized controlled trial. *J Natl Cancer Inst* 2006; **98:** 765–74.

10 Ronco G, Giorgi-Rossi P, Carozzi F, et al. Human papillomavirus testing and liquid-based cytology in primary screening of women younger than 35 years: results at recruitment for a randomised controlled trial. *Lancet Oncol* 2006; **7:** 547–55.

11 Carozzi F, Del Mistro A, Confortini M, et al. Reproducibility of HPV DNA testing by Hybrid Capture 2 in a screening setting: intralaboratory and interlaboratory quality control in seven laboratories participating in the same Clinical Trial. *Am J Clin Pathol* 2005; **124:** 1–6.

12 Luff RD. The Bethesda System for reporting cervical/vaginal cytologic diagnoses: a report of the 1991 Bethesda Workshop. *Hum Pathol* 1992; **23:** 719–21.

13 Carozzi F, Cecchini S, Confortini M, et al. Role of p16(INK4a) expression in identifying CIN2 or more severe lesions among HPV-positive patients referred for colposcopy after abnormal cytology. *Cancer* 2006; **108:** 119–23.

14 Dalla Palma P, Giorgi Rossi P, Collina G, et al. The risk of false-positive histology according to the reason for colposcopy referral in cervical cancer screening: a blind revision of all histological lesions found in the NTCC trial. *Am J Clin Pathol* 2008; **129:** 75–80.

15 Schatzkin A, Connor RJ, Taylor PR, Bunnag B. Comparing new and old screening tests when a reference procedure cannot be performed on all screenees. Example of automated cytometry for early detection of cervical cancer. *Am J Epidemiol* 1087; **125:** 672–78.

16 Spiegelhalter DJ, Thomas A, Best NG. WinBUGS version 1.2. Technical report. Cambridge, UK: MRC Biostatistics Unit, University of Cambridge, 1999.

17 Klaes R, Brenner A, Friedrich T, et al. p16INK4a immunochemistry improves interobserver agreement in the diagnosis of cervical intraepithelial neoplasia. *Am J Surg Pathol* 2002; **26:** 1389–99.

18 Agoff SN, Lin P, Morihara J, Mao C, Kiviat NB, Koutscky LA. P16(INK4a) expression correlates with degree of cervical neoplasia: a comparison with Ki-67 expression and detection of high-risk HPV types. *Mod Pathol* 2003; **16:** 665–73.

19 Dray M, Russel P, Dalrymple C, et al. p16(INK4a) as a complementary marker of high-grade intraepithelial lesions of the uterine cervix. I: experience with squamous lesions in 189 consecutive cervical biopsies. *Pathology* 2005; **37:** 112–24.

20 Hu L, Guo M, He Z, Thornton J, Mc Daniel LS, Hughson MD. Human papillomavirus genotyping and p16INK4a expression in cervical intraepithelial neoplasia of adolescents. *Mod Pathol* 2005; **18:** 267–73.

21 Wang SS, Trunk M, Schiffman M, et al. Validation of p16INK4a as a marker of oncogenic human papillomavirus infection in cervical biopsies from a population-based cohort in Costa Rica. *Cancer Epidemiol Biomarkers Prev* 2004; **13:** 1355–60.

22 Murphy N, Ring M, Killalea AG, et al. P16INK4A as marker for cervical diskaryosis: CIN and cGIN in cervical biopsies and ThinPrep smears. *J Clin Pathol* 2003; **56:** 56–63.

23 Nieh S, Chen SF, Chu TY, Lai HC, Fu E. Expression of p16 INK4A in Papanicolau smears containing atypical cells of undetermined significance from the uterine cervix. *Gynecol Oncol* 2003; **91:** 201–08.

24 Nieh S, Chen SF, Chu TY, Lai HC, Fu E. Expression of p16 INK4A in Papanicolau smears containing atypical glandular cells from the uterine cervix. *Acta Cytol* 2004; **48:** 173–80.

25 Guo M, Hu L, Baliga M, He Z, Hugson MD. The predictive value of p16(INK4a) and hybrid capture 2 human papillomavirus testing for high-grade cervocal intraepithelial lesions. *Am J Clin Pathol* 2004; **122:** 894–901.

26 Wenzensten N, Bergeron C, Cas F, Eschenbach D, Vinokurova S, von Knebel Doeberitz M. Evaluation of nuclear score for p16INK4a-stained cervical squamous cells in liquid-based cytology samples. *Cancer (Cancer Cytopathol)* 2005; **105:** 461–67.

27 Holladay EB, Logan S, Arnold J, Knesel B, Smith GD. A comparison of the clinical utility of p16(INK4a) immunolocalization with the presence of human papillomavirus by hybrid capture 2 for the detection of cervical dysplasia/neoplasia. *Cancer* 2006; **108:** 451–61.

28 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 $p16^{INK4a}$ positive cells to identify patients with high-grade cervical intraepithelial neoplasia. *Cancer* 2007; **111:** 58–66.

29 Meyer JL, Hanlon DW, Andersen BT, Rasmussen OF, Bisgaard K. Evaluation of $p16^{INK4a}$ expression in ThinPrep cervical specimens with the CINtec $p16^{INK4a}$ assay. *Cancer* 2007; **111:** 83–92.

30 Mao C, Balasubramanian A, Yu M, et al. Evaluation of a new p16(INK4A) ELISA test and high-risk HPV DNA test for cervical cancer screening: results from proof-of-concept study. *Int J Cancer* 2007; **120:** 2435–38.

31 Wang JL, Zheng BY, Li XD, Angstorm T, Lindstrom MS, Wallin KL. Predictive value of the alterations of p16INK4A, p14ARF, p53 and proliferating cell nuclear antigen expression in the progression of cervical cancer. *Clin Cancer Res* 2004; **10:** 2407–14.

32 Negri G, Vittadello F, Romano F, et al. p16INK4a expression and progression of low-grade intraepithelial neoplasia of the cervix. *Virchows Arch* 2004; **445:** 616–20.

33 Pretorius RG, Zhang WH, Belison JL, et al. Colposcopically directed biopsy, random cervical biopsy and endocervical curettage in the diagnosis of cervical intraepithelial neoplasia II or worse. *Am J Obstet Gynecol* 2004; **191:** 430–34.

34 Pretorius RG, Bao YP, Belison J, Burchette R, Smith J, Quiao YL. Inappropriate gold standard bias in cervical cancer screening studies. *Int J Cancer* 2007; **121:** 2218–24.