# 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

*Francesca Carozzi, Anna Gillio-Tos, Massimo Confortini, Annarosa Del Mistro, Cristina Sani, Laura De Marco, Salvatore Girlando, Stefano Rosso, Carlo Naldoni, Paolo Dalla Palma, Manuel Zorzi, Paolo Giorgi-Rossi, Nereo Segnan, Jack Cuzick, Guglielmo Ronco, and the NTCC working group*

*Lancet Oncol* 2013; 14: 168–76

Published Online December 21, 2012 http://dx.doi.org/10.1016/S1470-2045(12)70529-6

See Comment page 107

**Istituto per lo Studio e la Prevenzione Oncologica, Florence, Italy** (F Carozzi PhD, M Confortini PhD, C Sani MSc); **Cancer Epidemiology, Center for Experimental Research and Medical Studies, University of Turin, Italy** (A Gillio-Tos PhD, L De Marco PhD); **Veneto Oncology Institute, Istituto Di Ricovero e Cura a Carattere Scientifico, Padua, Italy** (A Del Mistro MD, M Zorzi MSc); **Santa Chiara Hospital, Trento, Italy** (S Girlando PhD, P Dalla Palma MD); **Centro per la Prevenzione Oncologica, Turin, Italy** (S Rosso MD, N Segnan MD, G Ronco MD); **Emilia-Romagna Region, Bologna, Italy** (C Naldoni MD); **Servizio Interaziendale Epidemiologia, Azienda Sanitaria Locale di Reggio Emilia, Reggio Emilia, Italy** (P Giorgi-Rossi PhD); **and Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, UK** (Prof J Cuzick PhD)

Correspondence to: Dr Guglielmo Ronco, Centro per la Prevenzione Oncologica, Via San Francesco da Paola 31, 10123 Torino, Italy **guglielmo.ronco@cpo.it**

## Summary

**Background** Immunostaining for p16-INK4A (henceforth p16) is a sensitive and specific method for detection of high-grade cervical intraepithelial neoplasia (CIN) in women infected with human papillomavirus (HPV), but longitudinal data have not been obtained. We investigated the relation between p16 status and risk of CIN during 3 years of follow-up.

**Methods** Women aged 25–60 years were enrolled between June 10, 2003, and Dec 31, 2004, in a multicentre randomised trial comparing HPV testing with cytology. HPV-positive women were referred for colposcopy and, in seven of nine centres, were tested for p16 overexpression by immunostaining. If no CIN was detected, these women were followed up at yearly intervals until clearance of HPV infection. The primary endpoint was histologically confirmed CIN of grade 2 or worse (CIN of grade 2 [CIN2], CIN of grade 3 [CIN3], or invasive cervical cancer) at recruitment or during follow-up. We calculated the absolute and relative risks by p16 status at recruitment. We also calculated the longitudinal sensitivity of p16 testing. Additionally, we assessed the relative sensitivity of an alternative strategy (referral to colposcopy and follow-up of only HPV-positive, p16-positive women) versus conventional cytology in two age groups. Percentages were weighted by the inverse of the tested fraction. The trial in which this study is nested is registered, number ISRCTN81678807.

**Findings** Of 1042 HPV-positive women who were tested for p16 with no CIN detected during the first round of screening, 944 (91%) had further HPV tests. 793 (84%) of these 944 were followed up until detection of CIN2 or worse, HPV infection clearance, or for at least 3 years. CIN2 or worse was detected during follow-up in more p16-positive women (31 of 365, 8·8% [95% CI 5·8–11·8]) than in p16-negative women (17 of 579, 3·7% [1·9–5·4]; relative risk [RR] 2·61 [95% CI 1·49–4·59]). RR was higher in women aged 35–60 years at recruitment (3·37 [1·39–8·15]) than in those aged 25–34 years (2·15 [1·00–4·61]), but age was not a significant modifier. CIN3 or worse was detected during follow-up in more p16-positive women (16 of 365, 4·4% [2·3–6·6]) than in p16-negative women (six of 579, 1·3% [0·2–2·3]; RR 3·90 [95% CI 1·57–9·68]). Longitudinal sensitivity of p16 testing for detection of CIN3 or worse during follow-up at all ages was 77·8% (95% CI 63·9–91·6). The relative sensitivity of the alternative strategy compared with conventional cytology was 2·08 (1·13–3·56) in women aged 35–60 years and 2·86 (1·28–5·36) in those aged 25–34 years. HPV-positive, p16-negative women aged 35–60 years had a higher cumulative risk of CIN3 or worse during recruitment or follow-up (2·0%, 95% CI 0·3–3·7) than did HPV-negative women (0·01%, 0–0·04) or those who were cytologically normal (0·04%, 0·02–0·09) at recruitment.

**Interpretation** p16 overexpression is a marker for CIN2 or worse or for development of CIN2 or worse within 3 years in HPV-positive women, especially those aged 35–60 years. HPV-positive, p16-positive women need immediate colposcopy and, if the assessment is negative, annual follow-up. Immediate colposcopy can be avoided in HPV-positive, p16-negative women, who can be safely managed with repeat screening after 2–3 year intervals.

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

## Introduction

Cross-sectional, two-sample studies and randomised controlled trials have shown that testing of human papillomavirus (HPV) DNA has greater sensitivity than cytology for detection of high-grade cervical intraepithelial neoplasia (CIN).[1,2] Randomised controlled trials have also established that screening based on HPV testing allows earlier diagnosis of persistent high-grade CIN[2] and is more effective in prevention of invasive cervical cancer.[2–5] However, HPV testing detects many transient spontaneously regressive infections, meaning that its specificity for high-grade CIN is low.[1,2] Therefore, methods are needed for selection of which HPV-positive women need colposcopy. In some randomised trials,[5,6] only women who had abnormal cytology or persistent HPV infections were referred to colposcopy. This

approach is effective, but entails short-term repeats of tests, which is disturbing for women[7] and leads to substantial loss to follow-up.[2] Therefore, approaches that do not necessitate short-term retesting would be useful.

In the New Technologies for Cervical Cancer screening (NTCC) study—a large randomised controlled trial comparing HPV DNA testing with cytology testing—HPV-positive women had postcolposcopy follow-up every year until HPV clearance.[4] Many new lesions were detected during such postcolposcopy clinical follow-up,[4] but this approach is expensive and distressing for women.[7] Identification of markers that would allow long intervals between testing for some HPV-positive women would therefore be useful.

p16-INK4A (henceforth p16) is a cyclin-dependent kinase inhibitor that is usually expressed at low concentration in healthy cells, but is overexpressed in cervical-cancer cell lines through mechanisms involving expression of the high-risk HPV E7 oncoprotein.[8,9] Therefore, p16 overexpression is an indicator of viral-induced deregulation of the cell cycle. Results of a nested substudy[10] of the NTCC trial showed that testing for p16 overexpression in HPV-positive women has high cross-sectional sensitivity and specificity for detection of high-grade CIN. Furthermore, the findings suggested that immediate colposcopy is not needed in HPV-positive women who do not overexpress p16 (henceforth p16 negative), allowing immediate referral rates to be reduced by 60%.

However, longitudinal data for the subsequent risk of high-grade CIN in women of different p16 status are needed to establish the appropriate frequency of re-testing in HPV-positive, p16-negative women, and in HPV-positive, p16-positive women with normal colposcopy results. We investigated the relation between p16 status at baseline and risk of CIN of grade 2 (CIN2) or worse during 3 years of follow-up. We specifically focused on detection of new CIN of grade 3 (CIN3) or worse, because the probability of progression to cancer from CIN3 within a short interval is higher than from CIN2. We also stratified data by age, as in all previous analyses of NTCC,[4,11–13] which showed differences between age groups for relative sensitivity, positive predictive value, and overdiagnosis of HPV testing compared with cytology.

## Methods

### Patients and procedures

The NTCC study was a randomised controlled trial with two preplanned recruitment phases in nine population-based cervical screening programmes in Italy.[4,11–13] Here, we report results of only women recruited during phase 2. Women aged 25–60 years who attended a new round of routine cervical-cancer screening between June 10, 2003, and Dec 31, 2004, were randomly assigned to receive conventional cytology (conventional group) or to standalone HPV-based screening (experimental group). Women were excluded from the study if they were a virgin, pregnant, had had a hysterectomy, or had been treated for CIN in the previous 5 years. Details about randomisation and masking have been reported previously.[13]

HPV DNA testing in the experimental group was done by Hybrid Capture 2 (Digene Corporation, Gaithersburg USA; now Qiagen, Hilden, Germany). Only probes designed to detect the high-risk HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68 were used.

During phase two, women in the experimental group were directly referred to colposcopy when their HPV test was positive with the conventional 1·0 pg/mL cutoff.[13] Specimens for p16 staining were obtained from these women at first colposcopy in seven of the nine study centres.[10] Some women having a colposcopy in one of these seven centres had no sample taken for p16 testing, mainly for reasons of organisation and because sample collection did not start immediately.[10] Additionally, a systematic (by colposcopy order) random sample of 20% of p16 specimens from women who had had no biopsy taken was discarded in five of these seven centres to reduce costs.[10] Overall, 519 (80%) of 647 women who had at least one colposcopy-directed biopsy at baseline and 651 (72%) of 900 without colposcopy-directed biopsy had a specimen taken for p16 testing.[10] Results were not used for management or treatment. We obtained multicentre and local research ethics approvals. Participants provided written informed consent.

### p16 immunostaining

Methods have been previously reported in detail.[10] Briefly, after preparation of one slide for cytology, we used 2 mL of residual PreservCyt fluid (Cytyc Corporation, Marlborough, MA, USA) with p16 samples obtained before colposcopy for a cytospin preparation. For immunostaining, we used the CINtec p16-INK4A Cytology kit (Dako Cytomation, now Roche mtm laboratories AG, Heidelberg, Germany), which applies a p16-specific monoclonal antibody (clone E6H4). We used haematoxylin as a counterstain. Cytospin preparation and staining were centralised. Positive control slides prepared from a cell pool containing residual clinical samples with high-grade squamous intraepithelial lesions and negative control slides from samples with negative cytological results were treated in the same way and used in each immunostaining procedure.

In 33 of 1170 women who had available p16 samples, the test was deemed to be unsatisfactory because the slides were obscured by blood or by inflammatory exudates or had insufficient numbers of cells (<500).[10] A p16-negative result was defined as no cell staining or staining of just morphologically normal endocervical, metaplastic, or atrophic cells. Staining of bacteria was not deemed to be a positive result. Staining of any other cells—eg, superficial, intermediate, and parabasal normal cells and all abnormal cells—was deemed to be a positive result. Indeed, in this HPV-positive population,

36% of samples with normal squamous epithelial cells stained for p16. Slides were independently read by FC and MC, who were masked to cytological and histological diagnosis. Discordant readings (41 [4%] of 1170) were resolved by consensus review.

### Postcolposcopy follow-up and endpoint assessment

As previously described,[4] women in the study (ie, in conventional and experimental groups) identified as having CIN2 or worse were treated and those with CIN of grade 1 (CIN1) were followed up with colposcopy. Women from both groups were recalled for repeat colposcopy, on the basis of colposcopy findings and results of cytology testing, according to routine local protocols. In the experimental group, HPV-positive women (who were to have had colposcopy) were actively recalled for annual repeats of HPV testing and liquid-based cytology until a negative HPV test was obtained. When follow-up liquid-based cytology identified atypical cells of undetermined significance or results that were more severe, women were referred for a repeat colposcopy. HPV-negative women in the experimental group and women in the conventional group who had normal initial cytology results were invited for a second screening about 3 years after the normal test. At the second screen, they were tested by conventional cytology and then managed according to the standard protocol of the centre.

The primary endpoint was histologically confirmed CIN2 or CIN3 or invasive cervical cancer (ie, CIN2 or worse) diagnosed within 3 years of colposcopy. We recorded test results and histological findings from the computerised registration systems of participating screening centres. For women who had a biopsy that was locally diagnosed as CIN1 or worse, all histological specimens were reviewed by a group of pathologists who were unaware of original diagnosis, randomisation, and p16 status.[4,10–14] They used morphological criteria[15] and no biomarkers. Adenocarcinoma in situ was grouped with CIN3. To obtain histological diagnoses that were made outside the trial, after the second round of screening, we linked the database of recruited women to those of the cancer registries (covering all centres except Viterbo) and of the pathology units present in the catchment areas of the NTCC study (histological

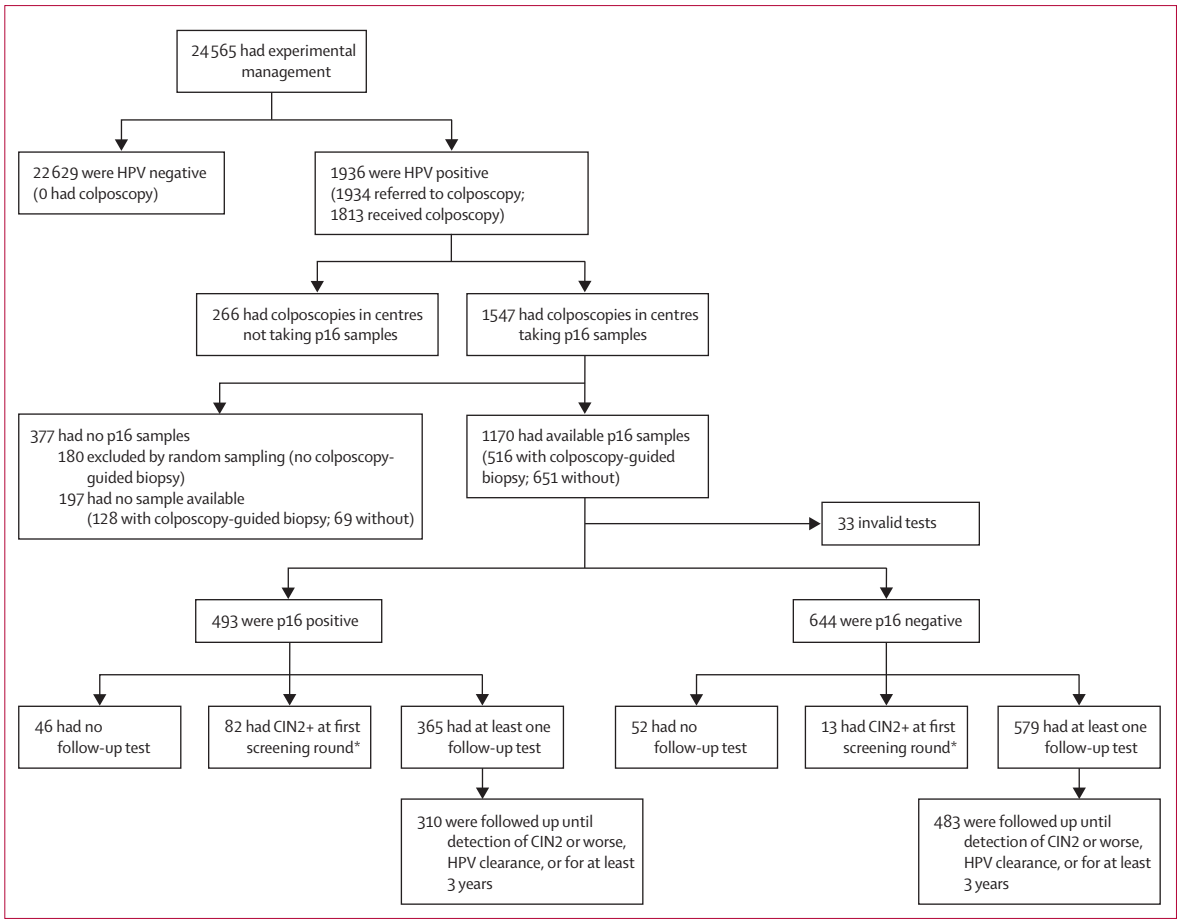


***Figure 1:*** **Phase two study profile**
*Because of differential inclusion of samples from colposcopies with and without colposcopy-guided biopsy, the crude data cannot be used to calculate absolute or relative risks.

diagnoses of CIN would probably have been registered in either the cancer registries or the pathology units).[4]

### Statistical analysis

Results for lesions detected within 6 months of initial colposcopy have been previously reported.[10] We computed the absolute and relative risks of a newly detected lesion during follow-up (ie, between 6 months and 3 years after colposcopy) by initial p16 result in HPV-positive women in the experimental group. We also calculated the longitudinal sensitivity of p16 testing—ie, the proportion of women who had CIN2 or worse diagnosed during follow-up who also had a positive result in initial p16 testing. Additionally, we computed the risks and longitudinal sensitivity for lesions detected either at recruitment or during follow-up. Because different proportions of women who did and did not have an initial biopsy had been tested for p16, most values (the risk of high-grade CIN in p16-positive and p16-negative women during postcolposcopy follow-up; longitudinal sensitivity; and the proportion of women who had had a colposcopy who were p16 negative, who provided a biopsy and did not have CIN2 or worse, and who had a colposcopy during follow-up) were obtained as weighted means of the corresponding proportions of women who had and who had not had a biopsy at baseline. We adjusted relative risks for initial biopsy.

We computed the absolute probabilities of detection of CIN3 or worse at the second screening round in women assigned to the experimental group in phase two who were HPV negative at baseline or who were assigned to the conventional group and were cytologically normal at baseline. Because these women had had no initial colposcopy and hence no postcolposcopy follow-up, detection of CIN3 or worse after 3 years provides an estimate of the cumulative incidence of persistent lesions that can be detected with cytology within 3 years of recruitment. We estimated the risk in women with negative cytology in the general population (mainly HPV negative), and we estimated the risk in p16-negative individuals who were HPV positive, therefore at higher previous risk. The purpose of the comparison was to assess whether p16 allows selection of a subpopulation in this a-priori, high-risk, HPV-positive population with a risk as low as that in cytologically normal women from the general population or in HPV-negative women.

Finally, we estimated the relative detection of CIN2 or worse at recruitment and during 3-year postcolposcopy follow-up if only HPV-positive, p16-positive women had been referred to colposcopy and postcolposcopy follow-up. We obtained this estimate—corresponding to relative sensitivity—by multiplying the relative detection in the same period recorded in phase 2 of the NTCC study with direct referral and follow-up of all HPV-positive women for the overall longitudinal sensitivity of p16 for lesions detected at recruitment or during postcolposcopy follow-up. Relative detection was estimated by intention to screen in phase two, including only lesions detected within 3 years of the first colposcopy.

Because we were not aware of a simple analytical solution for the variance of such a compound index (ie, the product of a sum of random variables for a ratio of other random variables), we used the Monte Carlo Markov Chain (MCMC) method.[16] We obtained its 95% CIs from the a-posteriori distribution, after sampling from beta distributions with parameters from the reported values (eg, positive cases over total tested cases of the pertinent group) of each proportion. We did the sampling with two MCMC runs with 10 000 cycles with WinBUGS (version 1.4.3).[17] We used the same approach to estimate the relative proportion of screened women who would have had new colposcopy during the clinical follow-up. For other analyses, we used SAS (version 8.2).

NTCC is registered, number ISRCTN81678807.

### Role of the funding source

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in this study; the corresponding author had full access to all the data and had final responsibility for the decision to submit for publication.

## Results

Results for p16 testing at recruitment have been reported previously.[10] Of the 1137 HPV-positive women tested for p16 at the first colposcopy, CIN2 or worse was detected within 6 months in 95 (8%; figure 1). Of the remaining 1042 women, 944 (91%) received further tests as part of the postcolposcopy follow-up (figure 1). Median duration of their follow-up was 1111 days (IQR 768–1366). 793 (84%)

| | p16-positive women | p16-negative women | Relative risk (95% CI)* | Longitudinal sensitivity of p16 (95% CI) |
|---|---|---|---|---|
| **CIN2 or worse** | | | | |
| All ages | 31/365 (8·8%, 5·8–11·8) | 17/579 (3·7%, 1·9–5·4) | 2·61 (1·49–4·59) | 66·9% (52·4–79·5) |
| Age 25–34 years | 13/151 (8·6%, 4·1–11·0) | 11/305 (4·1%, 1·7–6·4) | 2·15 (1·00–4·61) | 56·1% (35·5–76·7) |
| Age 35–60 years | 18/214 (8·3%, 4·6–11·9) | 6/274 (2·5%, 0·6–4·5) | 3·37 (1·39–8·15) | 76·9% (58·9–94·8) |
| **CIN3 or worse** | | | | |
| All ages | 16/365 (4·4%, 2·3–6·6) | 6/579 (1·3%, 0·2–2·3) | 3·90 (1·57–9·68) | 77·8% (63·9–91·6) |
| Age 25–34 years | 6/151 (3·9%, 0·9–7·0) | 4/305 (1·5%, 0·0–3·0) | 2·69 (0·79–9·20) | 61·8% (29·7–93·9) |
| Age 35–60 years | 10/214 (4·7%, 1·8–7·5) | 2/274 (0·8%, 0–1·9) | 6·05 (1·38–26·5) | 83·7% (62·8–100) |
| **CIN2** | | | | |
| All ages | 15/365 (4·4%, 2·2–6·5) | 11/579 (2·4%, 1·0–3·8) | 1·91 (0·90–4·08) | 58·9% (38·8–79·1) |
| Age 25–34 years | 7/151 (4·6%, 1·3–8·0) | 7/305 (2·6%, 0·7–4·5) | 1·84 (0·67–5·03) | 52·0% (25·3–78·8) |
| Age 35–60 years | 8/214 (3·6%, 1·2–6·0) | 4/274 (1·7%, 0·0–3·4) | 2·11 (0·65–6·81) | 66·7% (31·1–100) |

Data are number of women who had CIN detected during follow-up/number of women who had no CIN2 or worse detected at baseline but had follow-up (%, 95% CI) unless otherwise stated. Percentages are weighted by the inverse of the tested fraction, so do not correspond to the ratio of absolute numbers. All ages at recruitment. CIN=cervical intraepithelial neoplasia. CIN2=CIN of grade 2. CIN3=CIN of grade 3. *Adjusted for biopsy taken at initial colposcopy.

***Table 1:*** **Risk of detection of high-grade CIN during follow-up by baseline p16 immunostaining**

of the 944 women who had active follow-up were followed up until detection of CIN2 or worse or infection clearance, or for at least 3 years. 48 (5%) were actively followed up until detection of CIN2 or worse and 586 (62%) until a negative HPV test; 159 (17%) had a follow-up of at least 3 years and 151 (16%) had a shorter follow-up without disease or resolution of HPV infection (did not present for retesting).

Initial p16 result was significantly associated with the subsequent occurrence of CIN2 or worse during follow-up (table 1). The point estimate of relative risk for p16 was higher in women aged 35–60 years than in those aged 25–34 years (table 1), but the difference was not significant. The point estimate of longitudinal sensitivity was higher in the women aged 35–60 years than in those aged 25–34 years (table 1). The association between initial p16 result and the risk of CIN3 or worse during follow-up was particularly strong (table 1). Again, the point estimate of longitudinal sensitivity was higher in the older group than in the younger group (table 1). Risk of CIN2 was not associated with p16 results and longitudinal sensitivity of p16 testing for CIN2 was less than 70% at all ages (table 1).

During the first 3 years of postcolposcopy follow-up, 1023 (4·1%) of 24661 women in the experimental group had further colposcopy, compared with 253 (1·0%) of 24535 in the conventional group. Therefore, the proportion of enrolled women who underwent colposcopy during postcolposcopy follow-up was 4·02 times (95% CI 3·51–4·61) higher in the experimental group than in the conventional group. 386 (63·1%, 95% CI 59·5–66·6) of 617 colposcopies done during follow-up were for p16-negative women in the experimental group. Therefore, if clinical follow-up had been limited to HPV-positive, p16-positive women, the number of women in the experimental group who had further colposcopy would have been just 1·48 times (95% CI 1·21–1·71) higher than in the conventional group. Of 258 women in the experimental group who had a biopsy but did not have a diagnosis of CIN2 or worse during postcolposcopy follow-up, 160 (62·4%, 95% CI 56·2–68·7) were p16 negative, as were 45 (57·1%, 45·6–68·7) of 81 who had CIN1 detected. 36 (9·3%, 6·4–12·3) of 365 HPV-positive, p16-positive women had CIN1 detected during post-colposcopy follow-up.

When we assessed the risk of a CIN2, CIN3, or worse diagnosis at recruitment and during follow-up, we noted that it was higher in p16-positive women than in p16-negative women (table 2). Relative risk of CIN2 or worse during recruitment and follow-up in HPV-positive women was again higher in women who were p16 positive than in those who were p16 negative, with the greatest difference in the group aged 35–60 years (table 2).

HPV-positive, p16-negative women aged 35–60 years had a 2·0% (95% CI 0·3–3·7) risk of detection of CIN3 at recruitment or during follow-up (table 2). We detected no invasive cancer at recruitment or during follow-up, compared with three cases in HPV-positive, p16-positive women who were the same age (two at recruitment and one at follow-up). At the second round of screening (after about 3 years), we identified two (0·01%, 95% CI 0–0·04) cases of CIN3 and no invasive cancer in 16221 HPV-negative women aged 35–60 years, and seven cases of CIN3 or worse (including three invasive cancers; 0·04%, 95% CI 0·02–0·09) in 16940 women of the same age who were cytologically normal at recruitment.

If only HPV-positive, p16-positive women were referred to colposcopy and had postcolposcopy follow-up, the relative sensitivities of detection of CIN3 or worse and CIN2 would have been only slightly lower than when all HPV-positive women were referred to colposcopy and had post-colposcopy follow-up (table 3).

| | p16-positive women | p16-negative women | Relative risk (95% CI)* | Longitudinal sensitivity of p16 (95% CI) |
|---|---|---|---|---|
| **CIN2 or worse** | | | | |
| All ages | 113/493 (19·5%, 16·5–22·6) | 30/644 (5·2%, 3·4–7·0) | 3·74 (2·57–5·43) | 75·6% (63·5–87·7) |
| Age 25–34 years | 59/223 (22·5%, 17·7–27·3) | 21/349 (6·8%, 4·0–9·5) | 3·31 (2·10–5·22) | 68·9% (51·8–86·0) |
| Age 35–60 years | 54/270 (17·0%, 13·1–20·8) | 9/295 (3·5%, 1·3–5·8) | 4·81 (2·46–9·41) | 82·5% (66·4–98·6) |
| **CIN3 or worse** | | | | |
| All ages | 55/493 (9·7 %, 7·2–12·2) | 10/644 (1·7%, 0·7–2·8) | 5·57 (2·88–10·76) | 82·4% (67·8–97·0) |
| Age 25–34 years | 26/223 (9·9%, 6·3–13·4) | 5/349 (1·6%, 0·2–3·0) | 6·25 (2·41–16·22) | 76·4% (49·6–100·0) |
| Age 35–60 years | 29/270 (9·4%, 6·1–16·6) | 5/295 (2·0%, 0·3–3·7) | 4·72 (1·90–11·76) | 85·5% (68·4–100·0) |
| **CIN2** | | | | |
| All ages | 58/493 (9·9%, 7·5–12·2) | 20/644 (3·5%, 2·0–5·0) | 2·83 (1·74–4·61) | 68·2% (48·8–87·7) |
| Age 25–34 years | 33/223 (12·6%, 8·7–16·6) | 16/349 (5·2%, 2·7–7·6) | 2·41 (1·38–4·23) | 64·2% (42·7–85·7) |
| Age 35–60 years | 25/270 (7·6%, 5·1–10·1) | 4/295 (1·6%, 0·0–3·1) | 4·92 (1·71–14·09) | 74·6% (40·5–100·0) |

Data are number of women who had CIN detected at recruitment or during follow-up/number of HPV-positive women who were tested for p16 (%, 95% CI) unless otherwise stated. Percentages are weighted by the inverse of the tested fraction, so do not correspond to the ratio of absolute numbers. All ages at recruitment. CIN=cervical intraepithelial neoplasia. CIN2=CIN of grade 2. CIN3=CIN of grade 3. *Adjusted for biopsy taken at initial colposcopy.

*Table 2:* **Risk of detection of high-grade CIN at recruitment or during follow-up by baseline p16 immunostaining**

| | All HPV-positive women* | HPV-positive, p16-positive women |
|---|---|---|
| **CIN3 or worse** | | |
| Women aged 25–34 years | 3·74 (1·93–7·25) | 2·86 (1·28–5·36) |
| Women aged 35–60 years | 2·43 (1·46–4·04) | 2·08 (1·13–3·56) |
| **CIN2** | | |
| Women aged 25–34 years | 4·47 (2·51–7·97) | 3·05 (1·44–5·17) |
| Women aged 35–60 years | 1·92 (1·19–3·12) | 1·23 (0·60–2·48) |

Data in parentheses are 95% CI. Sensitivity calculated with data from recruitment and the 3-year follow-up period after the first colposcopy. Similar results considering only relative sensitivity at baseline have been reported previously.[10] HPV=human papillomavirus. CIN=cervical intraepithelial neoplasia. CIN2=CIN of grade 2. CIN3=CIN of grade 3. *Strategy actually applied in the trial; values differ from previously published data[10] because here they are truncated at 3 years after the first colposcopy.

***Table 3*: Estimated relative sensitivity of HPV testing versus cytology for histologically confirmed high-grade CIN with different strategies of referral and clinical follow-up**

## Discussion

We have shown that the result of an initial p16 test is a good predictor of CIN2, CIN3, or worse that become detectable by colposcopy in the subsequent 3 years in HPV-positive women, especially those aged 35–60 years. Prediction of CIN3 or worse was better than for CIN2, possibly because the time needed for progression from p16 overexpression to CIN2 is shorter than for progression to CIN3. Additionally, our findings suggest that a screening policy based on HPV testing with immediate referral to colposcopy and yearly postcolposcopy follow-up for HPV-positive, p16-positive women has a high relative sensitivity when compared with cytology-based screening.

In previously reported results,[10] p16 immunostaining was shown to have a cross-sectional sensitivity of 88% (95% CI 80–94) and a specificity of 61% (57–64) for CIN2 or worse in HPV-positive women. The relative sensitivity of HPV and p16 testing versus conventional cytology for women aged 35–60 years was 1·53 (1·15–2·02). These results suggested that only HPV-positive, p16-positive women need immediate colposcopy; this strategy resulted in a frequency of referral that was just 1·08 times (0·96–1·21) higher than with cytology.

In the NTCC study, all HPV-positive women were referred to colposcopy and, if no high-grade lesion was detected immediately, they were actively followed up at intervals of 1 year until clearance of HPV infection. Compared with cytology-based screening, this policy resulted in earlier detection of persistent lesions and fewer invasive cancers,[4] but entailed many colposcopies and test repeats. A strategy that limits immediate colposcopy and postcolposcopy follow-up to p16-positive women is also expected to confer high protection at a population level, because its relative sensitivity when compared with cytology is also high, but it would reduce the number of HPV-positive women who would have to have postcolposcopy follow-up visits, new colposcopies, and biopsies that would not lead to a diagnosis of CIN2 or worse. Because only CIN2 lesions or worse need treatment,[18] such biopsies do not have any advantages.

However, the cumulative risk of CIN3 was much higher in HPV-positive, p16-negative women (2·0%) than in HPV-negative women (0·01%). Although these results could be partly explained by the fact that some lesions were not yet detectable by cytology at round two or that some regressed,[19] these findings suggest that HPV-positive, p16-negative women should be retested earlier than HPV-negative women, who are normally tested after long intervals (≥5 years).

The rate of transition from high-grade CIN to cancer increases with time from occurrence of the lesion. In view of the low sensitivity of cytology, some CIN3 lesions first detected by HPV testing at baseline could have been present for a long time, which would mean that—if left untreated—the risk of progression to cancer might not be low, even in a short period. Therefore, both the longitudinal risk and the cross-sectional sensitivity (which establishes how many such lesions remain undetected) are relevant for the definition of safe intervals for retesting. The cross-sectional sensitivity of combinations of cytology and genotyping for HPV16 or HPV18 is less than 80%.[20] Women with negative results in these tests are recommended to repeat HPV testing after 1 year. Because the cross-sectional sensitivity of p16 is higher (91%),[10] increased intervals are reasonable.

In many European countries, cytologically normal women are retested after 3 years. The detection rate of CIN3 at the second round (roughly after 3 years) in women who were cytologically negative (and mostly HPV negative) at baseline (0·04%) was much lower than the risk of CIN3 during 3 years of follow-up in HPV-positive, p16-negative women (2·0%). Therefore, although these risks are not strictly comparable and although we did not record invasive cancers in HPV-positive, p16-negative women in 3 years, it seems safe to retest them for HPV infection at an interval shorter than 3 years and refer them to colposcopy in case of persistence (figure 2; panel). Conversely, in view of the low risk of new CIN3 (not likely to progress to invasive cervical cancer in a short period[25]), further postcolposcopy follow-up in HPV-positive, p16-negative women could be at increased intervals. The resulting savings depend on infection clearance in p16-negative women, which is, to our knowledge, unknown. In middle age, about 70–80% of prevalent HPV infections clear in 2–3 years.[26] This rate could be even higher in p16-negative women, because disruption of cell-cycle control resulting in p16 overexpression[8,9] is not present. Therefore, most p16-negative women—about 60% of HPV-positive women—would need just one repeat screen with no initial colposcopy or further follow-up.

A crucial issue for the management of HPV-positive, p16-negative women is the probability of progression of p16-negative CIN3 to cancer. If it was low, no detection of

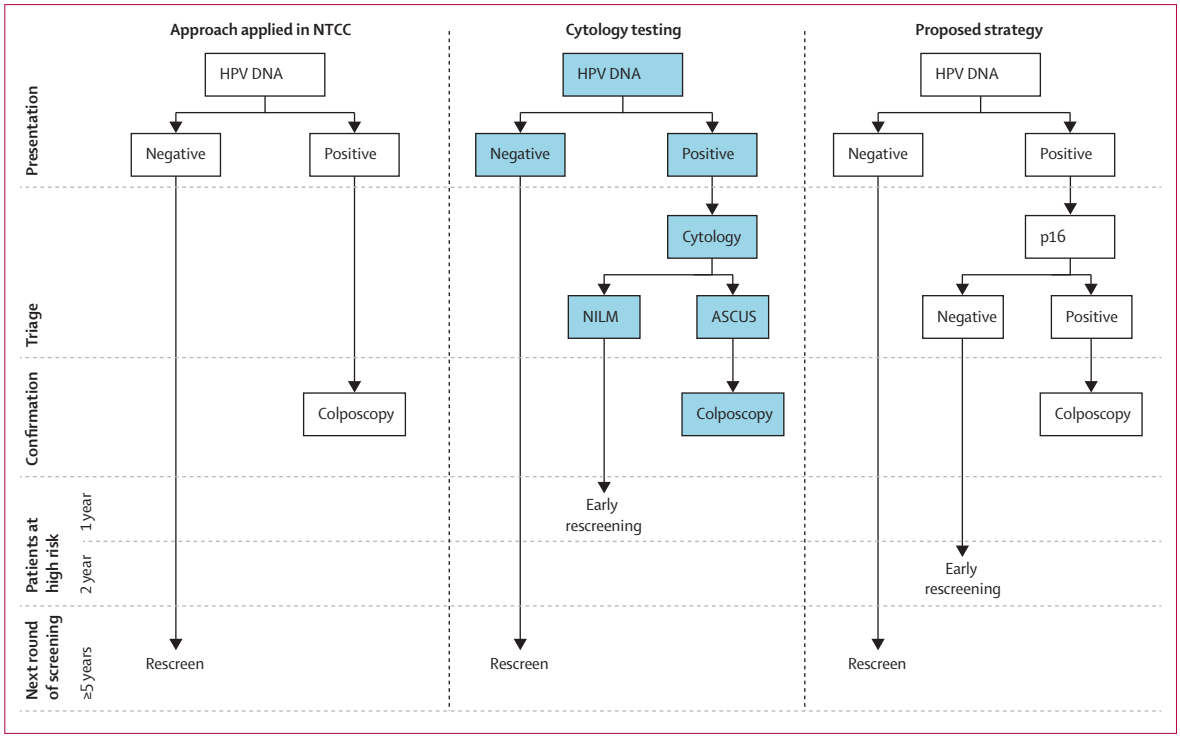


***Figure 2:*** **Management of HPV-positive women**
NTCC=New Technologies for Cervical Cancer Screening. HPV=human papillomavirus. NILM=negative for intraepithelial lesion or malignancy. ASCUS=atypical cells of undetermined significance.

CIN3 would be advantageous and long intervals between retests would be advisable. Some evidence supports this strategy,[21–24,27] but is not conclusive.

Prevalent infections are usually more recent in younger than in older women,[26] which could explain differences by age in point estimates of p16 longitudinal accuracy. The point estimate of p16 specificity at baseline was lower in women older than 35 years than in those younger than 35 years in a previous report.[10] Therefore, HPV infection had more frequently already progressed to p16 overexpression at recruitment in older than in younger women, although not to high-grade CIN. Application of HPV-based screening in young women is limited by overdiagnosis of regressive CIN2.[4] However, at any age, the longitudinal accuracy of p16 testing could decrease between the first and subsequent HPV screening rounds if intervals are too short because newly detected infections would at most have been present since the previous screen and would therefore have been recently acquired.

The NTCC study was population based and was nested in routine organised screening in a low-risk population. More than 70% of eligible women were enrolled,[4] suggesting that results are applicable to routine practice. The completeness of clinical follow-up was high and was equal in p16-positive and p16-negative women. In the main NTCC trial,[4] the proportion of women with CIN2 or worse at the second round of screening was similar for those who did and did not attend organised programmes, suggesting that completeness of data was high.

Notably, we used cytospin preparation instead of liquid-based cytology, which is usually used. Good-quality immunocytochemical staining can be achieved with cytology slides prepared and fixed in different ways. The UK National External Quality Assessment Service for Immunocytochemistry cytology module deemed that the highest quality of immunocytochemical staining on in-house control slides was achieved with cell block sections followed by cytospins, formalin-fixed paraffin-embedded tissue sections, liquid-based-cytology slides, and conventional smears.[28] Indeed, the previously reported cross-sectional sensitivity of p16 immunostaining of 88%[10] was similar to what has been recorded with liquid-based cytology in women with abnormal cytology, ranging from 78%[29] to 96%.[30]

We used some morphological criteria in addition to p16 staining. Double staining for both p16 and Ki-67 was introduced in Europe in March, 2010. This change was intended to avoid the need for morphological criteria and to increase specificity and reproducibility. The assay that we applied is no longer commercially available, although assays that are based on different antibodies for p16 are. Our results do not strictly apply to double staining. The antibody for p16 that we used is the same as is used with double staining, but different sensitivity and specificity

### *Panel:* Research in context

#### Systematic review

We searched Medline and PubMed in 2005, and again in 2012. Overall, we included reports published before Sept 30, 2012, that were identified with the keywords "p16" and "cervical screening". We identified 39 relevant reports, most of which were cross-sectional studies of the value of p16 testing for detection of disease, especially its clinical potential as a triage test for borderline and low-grade cytological results. Four longitudinal studies[21–24] showed a promising prognostic value of p16 for progression or recurrence after treatment, but were all based on staining of sections from biopsies of histologically confirmed cervical intraepithelial neoplasia (CIN) lesions. We identified no longitudinal study of the use of p16 staining of cytological material.

#### Interpretation

As far as we are aware, ours is the first study of p16 testing as a triage test after primary screening that is based on infection with human papillomavirus (HPV) with follow-up assessment. HPV screening programmes will soon be implemented in many countries. For women aged 35 years or older who are HPV positive (3–5%), p16 testing could be used as triage test to identify individuals at risk of developing CIN of grade 3 within the next 3 years. These women need immediate colposcopy and, if the assessment is negative, annual follow-up. Immediate colposcopy can be avoided in HPV-positive, p16-negative women, who can be safely managed with repeat screening after 2–3 years.

could be caused by the use of Ki-67 counterstaining instead of morphological criteria. In studies done in women with atypical cells of undetermined significance or low-grade squamous intraepithelial lesions,[31] in a colposcopy referral population,[32] and in HPV-positive women who were cytologically normal,[33] double staining had similar cross-sectional accuracy to what we reported in HPV-positive women with p16 alone.[10] Longitudinal accuracy depends mainly on the natural history of HPV-infected cells that do or do not overexpress p16—ie, on their rates of progression to high-grade CIN. Therefore, longitudinal accuracy of double staining would be expected to be similar to that of staining for p16 alone. Similarly, cross-sectional accuracy is deemed to be sufficient for validation of HPV DNA tests for screening because its longitudinal accuracy mainly depends on the transition rates from HPV infection to CIN.[34] Double-testing studies comparing double staining to stand-alone p16 staining, with the design recommended for HPV DNA assays' validation[34] are needed. Independent confirmatory studies on the longitudinal accuracy of p16—preferably with double staining with Ki-67—are advisable.

#### Contributors

GR was the NTCC project leader and designed the study with FC, AG-T and ADM. AG-T, ADM, LDM, and SG provided the cervical-cell samples. MC, CN, PDP, MZ, PG-R, and NS organised the fieldwork, including the follow-up. CS did the immunostaining procedure under the supervision of FC. MC and FC assessed the immunostained slides. SR and GR did the statistical analysis. JC and GR drafted the report. All authors critically revised the report.

#### Conflicts of interest

We declare that we have no conflicts of interest.

#### Acknowledgments

The NTCC main trial was financially supported by the European Union (Europe against cancer contracts S1.2.327046 and SPC2002473); the Italian Ministry of Health (special project Valutazione di nuove tecnologie per lo screening del cervicocarcinoma and L138/2004 project NTCC: Nuove tecnologie per lo screening del cervicocarcinoma); the Regional Health administrations of Piemonte, Tuscany, Veneto, and Emilia Romagna; and by the Public Health Region of Lazio Region. The nested p16 study and follow-up was supported by the Italian Ministry of Health (Special Project Valutazione di nuove tecnologie per lo screening del cervicocarcinoma—follow up) and by the European Union (PREHDICT project FP7 grant agreement No 242061). We thank all the staff for assistance in the running of the study and the thousands of women who participated.

#### References

1 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.

2 Ronco G, Meijer CJLM. HPV screening: available data and recommendations for clinical practice. *Curr Cancer Ther Rev* 2010; **6:** 104–09.

3 Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. *N Engl J Med* 2009; **360:** 1385–94.

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:** 249–57.

5 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.

6 Naucler P, Ryd W, Tornberg S, et al. Human papillomavirus and Papanicolau tests to screen for cervical cancer. *N Engl J Med* 2007; **357:** 1589–97.

7 Waller J, McCaffery K, Kitchener H, Nazroo J, Wardle J. Women's experiences of repeated HPV testing in the context of cervical cancer screening: a qualitative study. *Psychooncology* 2007; **16:** 196–204.

8 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.

9 McLaughlin-Drubin ME, Crum CP, Münger K. Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. *Proc Natl Acad Sci USA* 2011; **108:** 2130–35.

10 Carozzi F, Confortini M, Dalla Palma P, et al. 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:** 937–45.

11 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.

12 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.

13 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.

14 Dalla Palma P, Giorgi Rossi P, Collina G, et al. The reproducibility of CIN diagnoses among different pathologists: data from histology reviews from a multicentre randomized study. *Am J Clin Pathol* 2009; **132:** 125–32.

15 Richart RM. Cervical intraepithelial neoplasia. *Pathol Annu* 1973; **8:** 301–23.

16 Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Satistical Science* 1992; **7:** 457–511.

17 Spiegelhalter DJ, Thomas A, Best NG. WinBUGS: technical Report. Cambridge: MRC Biostatistics Unit, University of Cambridge, 1999.

18 Arbyn M, Anttila A, Jordan J, et al, eds. European guidelines for quality assurance on cervical cancer screening, 2nd edn. Luxembourg: Office for Official Publications of the European Communities, 2008.

19 Ronco G, Segnan N. HPV testing for primary cervical cancer screening. *Lancet* 2007; **370:** 1740–42.

20 Castle PE, Stoler MH, Wright TC, Sharma A, Wright TL, Behrens CM. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. *Lancet Oncol* 2011; **12:** 880–90.

21 Ozaki S, Zen Y, Inoue M. Biomarker expression in cervical intraepithelial neoplasia: potential progression predictive factors for low-grade lesions. *Hum Pathol* 2011; **42:** 1007–12.

22 del Pino M, Garcia S, Fusté V, et al. Value of p16(INK4a) as a marker of progression/regression in cervical intraepithelial neoplasia grade 1. *Am J Obstet Gynecol* 2009; **201:** 488.

23 Anschau F, Schmitt VM, Lambert AP, Gonçalves MA, Machado DC. Transition of cervical carcinoma in situ to invasive cancer: role of p16 INK4a expression in progression and in recurrence. *Exp Mol Pathol* 2009; **86:** 46–50.

24 Hariri J, Øster A. The negative predictive value of p16INK4a to assess the outcome of cervical intraepithelial neoplasia 1 in the uterine cervix. *Int J Gynecol Pathol* 2007; **26:** 223–28.

25 Plummer M, Peto J, Franceschi S, for the International Collaboration of Epidemiological Studies of Cervical Cancer. Time since first sexual intercourse and the risk of cervical cancer. *Int J Cancer* 2012; **130:** 2638–44.

26 Rodríguez AC, Schiffman M, Herrero R, et al. Longitudinal study of human papillomavirus persistence and cervical intraepithelial neoplasia grade 2/3: critical role of duration of infection. *J Natl Cancer Inst* 2010; **102:** 315–24.

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

28 Kirbis IS, Maxwell P, Fležar MS, Miller K, Ibrahim M. External quality control for immunocytochemistry on cytology samples: a review of UK NEQAS ICC (cytology module) results. *Cytopathology* 2011; **22:** 230–37.

29 Holladay EB, Logan S, Arnold J, Knesek 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 dysplasias/neoplasias. *Cancer* 2006; **108:** 451–61.

30 Wentzensen NM, 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 women with cervical intraepithelial neoplasia. *Cancer* 2007; **111:** 58–66.

31 Schmidt D, Bergeron C, Denton KJ, Ridder R. p16/Ki-67 dual stain cytology in the triage of ASCUS and LSIL Papanicolau cytology: results from the European equivocal or mildly abnormal Papanicolau cytology study. *Cancer Cytopathol* 2011; **119:** 158–66.

32 Wentzensen N, Schwartz L, Zuna RE, et al. Performance of p16/Ki-67 immunostaining to detect cervical cancer precursors in a colposocpy referral population. *Clin Cancer Res* 2012; **18:** 4154-62.

33 Petry KU, Schmidt D, Scherbring S, et al. Triaging Pap cytology negative, HPV positive cervical cancer screening results with p16/Ki-67 Dual-stained cytology. *Gynecol Oncol* 2011; **121:** 505–09.

34 Meijer CJ, Berkhof J, Castle PE, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *Int J Cancer* 2009; **124:** 516–20.