IJC
International Journal of Cancer

Combined use of cytology, p16 immunostaining and genotyping for triage of women positive for high-risk human papillomavirus at primary screening

Jack Cuzick[1], Rachael Adcock[1], Francesca Carozzi[2], Anna Gillio-Tos[3], Laura De Marco[3], Annarosa Del Mistro[4], Helena Frayle[4], Salvatore Girlando[5], Cristina Sani[2], Massimo Confortini[2], Manuel Zorzi[6], Paolo Giorgi-Rossi[7], Raffaella Rizzolo[8], and Guglielmo Ronco[8,9], On behalf of the New Technologies for Cervical Cancer Screening Working Group

[1]Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, United Kingdom
[2]Regional Cancer Prevention Laboratory, Oncological Network, Prevention and Research Institute (ISPRO), Florence, Italy
[3]Cancer Epidemiology Unit, Department of Medical Sciences, C.E.R.M.S, University of Turin, Turin, Italy
[4]Immunology and Diagnostic Molecular Oncology Unit, Veneto Institute of Oncology IOV IRCCS, Padua, Italy
[5]Pathology Unit, S. Chiara Hospital, Trento, Italy
[6]Veneto Tumour Registry, Azienda Zero, Padua, Italy
[7]Epidemiology Unit, Azienda USL, IRCCS di Reggio Emilia, Reggio Emilia, Italy
[8]Centre for Cancer Prevention, AOU Città della Salute e della Scienza, Torino, Italy
[9]International Agency for Research on Cancer (IARC), Lyon, France

Human papillomavirus (HPV) testing is very sensitive for primary cervical screening but has low specificity. Triage tests that improve specificity but maintain high sensitivity are needed. Women enrolled in the experimental arm of Phase 2 of the New Technologies for Cervical Cancer randomized controlled cervical screening trial were tested for high-risk HPV (hrHPV) and referred to colposcopy if positive. hrHPV-positive women also had HPV genotyping (by polymerase chain reaction with GP5+/GP6+ primers and reverse line blotting), immunostaining for p16 overexpression and cytology. We computed sensitivity, specificity and positive predictive value (PPV) for different combinations of tests and determined potential hierarchical ordering of triage tests. A number of 1,091 HPV-positive women had valid tests for cytology, p16 and genotyping. Ninety-two of them had cervical intraepithelial neoplasia grade 2+ (CIN2+) histology and 40 of them had CIN grade 3+ (CIN3+) histology. The PPV for CIN2+ was >10% in hrHPV-positive women with positive high-grade squamous intraepithelial lesion (61.3%), positive low-grade squamous intraepithelial lesion (LSIL+) (18.3%) and positive atypical squamous cells of undetermined significance (14.8%) cytology, p16 positive (16.7%) and, hierarchically, for infections by HPV33, 16, 35, 59, 31 and 52 (in decreasing order). Referral of women positive for either p16 or LSIL+ cytology had 97.8% sensitivity for CIN2+ and women negative for both of these had a 3-year CIN3+ risk of 0.2%. Similar results were seen for women being either p16 or HPV16/33 positive. hrHPV-positive women who were negative for p16 and cytology (LSIL threshold) had a very low CIN3+ rate in the following 3 years. Recalling them after that interval and referring those positive for either test to immediate colposcopy seem to be an efficient triage strategy. The same applies to p16 and HPV16.

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

Key words: cervical screening, HPV testing, triage of HPV-positive women, cytology, p16 immunocytochemistry, HPV genotyping

Abbreviations: ASC-US: atypical squamous cells of undetermined significance; CIN: cervical intraepithelial neoplasia; CIN3+: CIN grade 3+; CIN2+: CIN grade 2+; HC2: Hybrid capture 2; HPV: human papilloma virus; hrHPV: high-risk HPV; HSIL: high-grade squamous intraepithelial lesion; LBC: liquid-based cytology; LSIL: low-grade squamous intraepithelial lesion; NPV: negative predictive value; NTCC: New Technologies for Cervical Cancer; PCR: polymerase chain reaction; PPV: positive predictive value; RCT: randomized controlled trial; STM: standard transport medium

DOI: 10.1002/ijc.32973
History: Received 18 Nov 2019; Accepted 28 Feb 2020; Online 14 Mar 2020

Correspondence to: Raffaella Rizzolo, E-mail: raffaella.rizzolo@cpo.it

What's new?

Human papillomavirus (HPV) testing has revolutionized cervical cancer screening but alone cannot accurately predict the risk of cancerous lesions. Here the authors assessed its combined use with other predictive exams, cytology and p16 staining. They show that women testing positive for high-risk HPV but negative for p16 and either cytology or HPV16 have a very low rate of cervical intraepithelial neoplasia or worse (CIN3+) in the three years after the exam and can be safely recalled after that time period, reducing overall colposcopy referrals.

Introduction

Screening based on human papilloma virus (HPV) testing allows earlier diagnosis of high-grade cervical intraepithelial lesions (cervical intraepithelial neoplasia [CIN]) than cytology-based screening and is more effective in preventing subsequent invasive cervical cancer.[1,2] However, the specificity of HPV testing for high-grade CIN is lower,[3] and better methods are needed for selecting which HPV-positive women need immediate colposcopy. Randomized controlled trials (RCTs) have shown that referring to colposcopy only those HPV-positive women who also had abnormal cytology or persistent HPV infection leads to increased efficacy *versus* cytology-based screening, without increasing the biopsy rate.[2] However, short-term repeat tests are needed, which can produce anxiety[4] and entail appreciable loss to follow-up.[5]

Triage protocols that can safely allow longer intervals for low-risk women are desirable. For women known to be positive for high-risk HPV (hrHPV), we previously found p16 overexpression,[6] abnormal cytology (atypical squamous cells of undetermined significance [ASC-US] or higher),[7] and infection by HPV16, 33 or 35 types[8] to have cross-sectional sensitivities for CIN grade 3+ (CIN3+) of 91, 88 and 67, respectively. Here we consider the safety of restricting immediate colposcopy to those hrHPV-positive women who are positive for one or more of these tests, while returning the remainder to 3-year follow-up.

For this purpose, we used the material collected in Phase 2 of the New Technologies for Cervical Cancer (NTCC) screening RCT where all HPV-positive women were referred to colposcopy thus avoiding verification bias. Given the long interval needed for progression from <CIN3 to invasive cancer,[9] CIN3 present at study entry and left untreated is the primary concern for cancer development in the next 3 years. However, as some CIN3 present at baseline could have been missed by the first colposcopy, we also considered new lesions detected within 3 years in all HPV-positive women.

Methods

NTCC is a randomized trial conducted within nine population-based cervical screening programs in Italy. Women aged 25–60 years who were not pregnant, had never undergone hysterectomy, had not been treated for CIN in the last 5 years and who were attending for a new routine cervical screening appointment were randomly assigned to conventional cytology or to HPV-based screening, either in combination with liquid-based cytology (Phase 1)[10,11] or alone (Phase 2).[12]

During Phase 2, hrHPV testing was done using Hybrid capture 2 (HC2; Qiagen, Hilden, Germany) on samples of cervical cells collected in standard transport medium (STM; Qiagen). Women were referred for colposcopy if the HPV test was positive.[12] As a rule, women with CIN grade 2+ (CIN2+) were treated and those <CIN2 followed up colposcopically. If no CIN was detected, hrHPV-positive women were recalled for annual repeat testing with HC2 and ThinPrep liquid-based cytology (LBC) for as long as their HPV test remained positive and referred to colposcopy only if cytology became ASC-US or higher. Women from both arms who were screen negative at baseline were invited for a second screening round 3 years later using conventional cytology and managed according to the standard protocol for each center. Full details have been reported previously.[1]

NTCC is registered as an International Standard Randomized Controlled Trial, number ISRCTN81678807. We obtained multicenter and local research ethics approvals and written informed consents from all women.

Genotyping

During Phase 2, in all centers except Verona, residual material after a positive HC2 test was stored as 400 μl aliquots in STM at −80°C. Only the first HC2-positive sample from each woman was considered in this analysis. Genotyping was performed blind to histology results by polymerase chain reaction (PCR) with GP5+/GP6+ primers, followed by reverse line blot genotyping assay.[13] Analysis of HPV genotyping was restricted to the 13 hrHPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68) detected by HC2.

Cytology and p16 immunostaining

In the experimental arm of Phase 2, ThinPrep cytology was routinely prepared at first colposcopy. In all study centers except Verona and Viterbo, after preparation of one slide for cytology, 2 ml of the residual fluid was shipped for centralized immunostaining of cytospin slides with the CINtec™ p16-INK4A cytology kit.[6] At the time of testing, the dual stain including ki-67 was not available. For logistical reasons, sample collection started at different times in different centers. In five centers, a random sample of 20% of specimens from women who had no biopsy taken at colposcopy was discarded to reduce costs. Methods have been described previously in detail.[7]

A p16-INK4A-negative result was defined as no cell staining or staining of just morphologically normal endocervical, metaplastic or atrophic cells or bacteria. The presence of any other p16-positive cell, including superficial, intermediate and parabasal normal and all abnormal cells, was defined as positive. Slides were independently read by two investigators, blind to cytological and histological diagnosis, and discordant readings were resolved by consensus review blinded to all other data (except HPV positivity).

Endpoint assessment

Endpoints were histologically confirmed CIN3+ and CIN2+, which included invasive cervical cancer and adenocarcinoma *in situ*. At the end of the recruitment phase, all histological specimens taken within 1 year of referral to colposcopy and locally diagnosed as CIN1 or higher were reviewed by a group of pathologists who were blinded to the original diagnosis and randomization. Centrally reviewed biopsies after any follow-up colposcopy within 3.5 years of entry were also obtained (to allow for a small delay in attendance) and included in a "3-year" longitudinal analysis. Random biopsies were not taken in women when no abnormal area was seen on colposcopy.

Statistical analysis

The cohort consisted of all hrHPV-positive women who had valid tests for cytology, p16 immunostaining and genotyping. A "baseline" analysis was based on disease detected within 12 months of the first referral to colposcopy. We also considered disease detected within 3.5 years after the first colposcopy to measure disease up to and including the usual time until the next regular screening. This is denoted as "detected within 3 year follow-up" below. Three and a half year follow-up was complete for 84.3% of HPV+ women without disease (CIN2+) detected at baseline.

For the analysis, we considered five cytology categories (unsatisfactory, normal, ASC-US, low-grade squamous intraepithelial lesion [LSIL], positive high-grade squamous intraepithelial lesion [HSIL+]) and two for p16 (positive/negative).

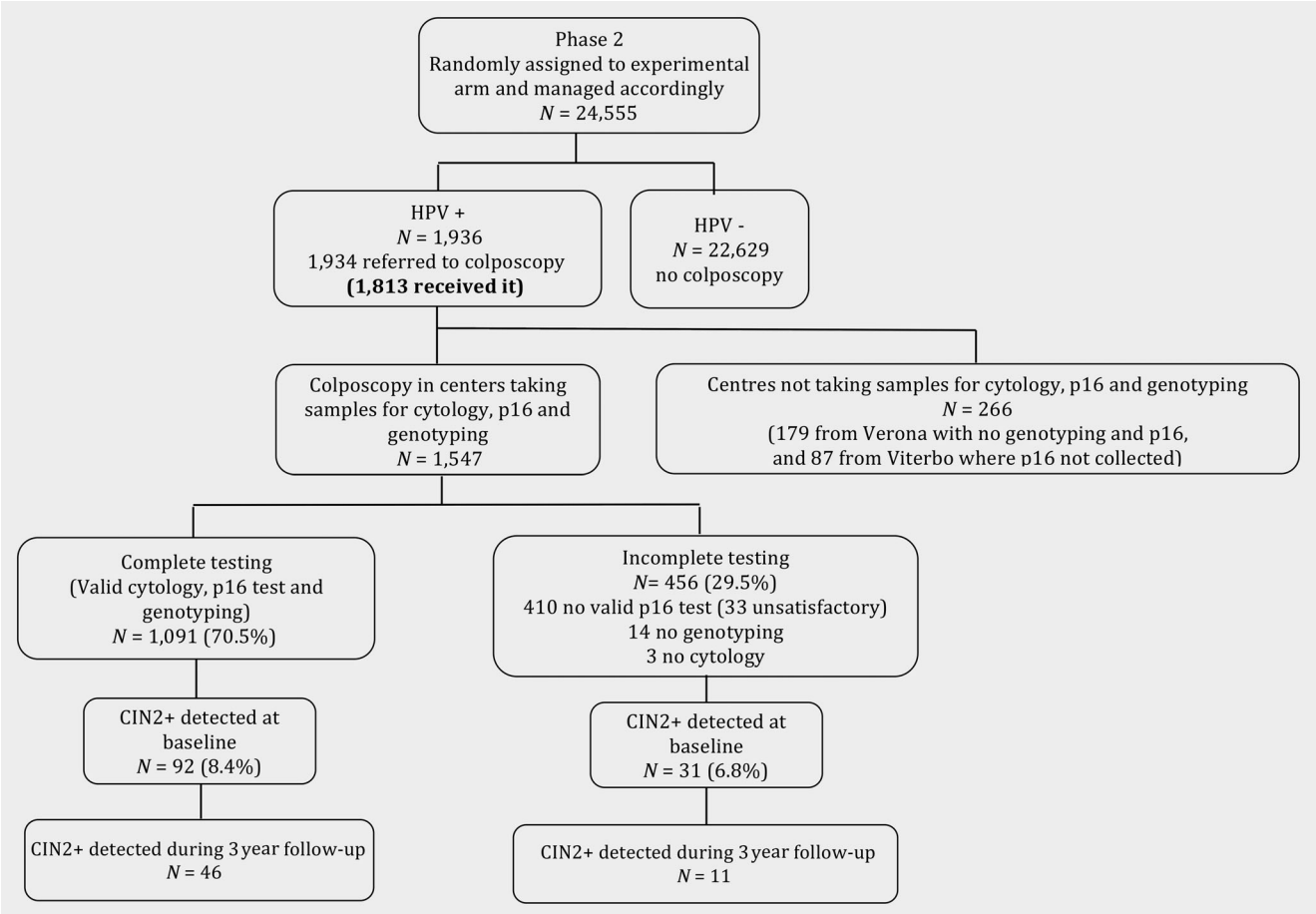

Figure 1. Sample selection process.

Table 1. Accuracy for detection of CIN2+ and CIN3+ detected at baseline[1]

(*a*) Combinations of HSIL+ cytology, p16 immunostaining, and HPV genotypes for HPV-positive women (*n* = 1,091)						
	HSIL+	**HSIL+ or p16+**	**HSIL+ or HPV16+**	**HSIL+ or HPV16/33+**	**HSIL+ or p16+ or HPV16+**	**HSIL+ or p16+ or HPV16/33+**
N positive (%)	80 (7.33)	474 (43.45)	342 (31.35)	367 (33.64)	631 (57.84)	646 (59.21)
			CIN2+ (*n* = 92)			
N	49	82	65	69	86	88
Sensitivity	53.26 (42.56, 63.74)	89.13 (80.92, 94.66)	70.65 (60.24, 79.69)	75.00 (64.89, 83.45)	93.48 (86.34, 97.57)	95.65 (89.24, 98.80)
Specificity	96.90 (95.62, 97.88)	60.76 (57.65, 63.80)	72.27 (69.38, 75.03)	70.17 (67.23, 72.99)	45.45 (42.33, 48.59)	44.14 (41.04, 47.29)
PPV	61.25 (49.70, 71.94)	17.30 (14.00, 21.01)	19.01 (14.99, 23.57)	18.80 (14.93, 23.18)	13.63 (11.05, 16.56)	13.62 (11.07, 16.51)
1-NPV	4.25 (3.09, 5.69)	1.62 (0.78, 2.96)	3.60 (2.39, 5.20)	3.18 (2.02, 4.73)	1.30 (0.48, 2.82)	0.90 (0.25, 2.29)
			CIN3+ (*n* = 40)			
N	24	38	32	34	38	39
Sensitivity	60.00 (43.33, 75.14)	95.00 (83.08, 99.39)	80.00 (64.35, 90.95)	85.00 (70.16, 94.29)	95.00 (83.08, 99.39)	97.50 (86.84, 99.94)
PPV	30.00 (20.26, 41.28)	8.02 (5.74, 10.84)	9.36 (6.49, 12.95)	9.26 (6.50, 12.71)	6.02 (4.30, 8.17)	6.04 (4.33, 8.16)
1-NPV	1.58 (0.91, 2.56)	0.32 (0.04, 1.17)	1.07 (0.46, 2.09)	0.83 (0.30, 1.80)	0.43 (0.05, 1.56)	0.22 (0.01, 1.25)

(*b*) Combinations of LSIL+ cytology, p16 immunostaining and HPV genotypes for HPV-positive women (*n* = 1,091)										
	LSIL+	**HPV16+**	**HPV16/33+**	**p16+**	**LSIL+ or p16+**	**LSIL+ or HPV16+**	**LSIL+ or HPV16/33+**	**p16+ or HPV16+**	**LSIL+ or p16+ or HPV16+**	**LSIL+ or p16+ or HPV16/33+**
N positive (%)	372 (34.10)	308 (28.23)	397 (36.39)	468 (42.90)	647 (59.30)	564 (51.70)	581 (53.25%)	629 (57.65)	761 (69.75)	771 (70.67)
					CIN2+ (*n* = 92)					
N	68	50	57	78	90	78	79	85	90	91
Sensitivity	73.91 (63.71, 82.52)	54.35 (43.63, 64.78)	61.96 (51.24, 71.88)	84.78 (75.79, 91.42)	97.83 (92.37, 99.74)	84.78 (75.79, 91.42)	85.87 (77.05, 92.26)	92.39 (84.95, 96.89)	97.83 (92.37, 99.74)	98.91 (94.09, 99.97)
Specificity	69.57 (66.61, 72.41)	74.17 (71.34, 76.86)	65.97 (62.93, 68.90)	60.96 (57.86, 64.00)	44.24 (41.14, 47.39)	51.35 (48.20, 54.49)	49.75 (46.60, 52.90)	45.55 (42.42, 48.69)	32.83 (29.92, 35.84)	31.93 (29.05, 34.92)
PPV	18.28 (14.48, 22.59)	16.23 (12.30, 20.84)	14.36 (11.06, 18.20)	16.67 (13.40, 20.36)	13.91 (11.34, 16.82)	13.83 (11.09, 16.96)	13.60 (10.92, 16.66)	13.51 (10.94, 16.44)	11.83 (9.62, 14.34)	11.80 (9.61, 14.29)
1-NPV	3.34 (2.15, 4.93)	5.36 (3.89, 7.18)	5.04 (3.54, 6.94)	2.25 (1.23, 3.74)	0.45 (0.05, 1.62)	2.66 (1.46, 4.42)	2.55 (1.36, 4.32)	1.52 (0.61, 3.10)	0.61 (0.07, 2.17)	0.31 (0.01, 1.73)
					CIN3+ (*n* = 40)					
N	28	26	28	36	39	35	36	37	39	40
Sensitivity	70.00 (53.47, 83.44)	65.00 (48.32, 79.37)	70.00 (53.47, 83.44)	90.00 (76.34, 97.21)	97.50 (86.84, 99.94)	87.50 (73.20, 95.81)	90.00 (76.34, 97.21)	92.50 (79.61, 98.43)	97.50 (86.84, 99.94)	100.00 (91.19, 100.00)

(Continues)

Table 1. Accuracy for detection of CIN2+ and CIN3+ detected at baseline (Continued)

(b) Combinations of LSIL+ cytology, p16 immunostaining and HPV genotypes for HPV-positive women (n = 1,091)										
	LSIL+	HPV16+	HPV16/33+	p16+	LSIL+ or p16+	LSIL+ or HPV16+	LSIL+ or HPV16/33+	p16+ or HPV16+	LSIL+ or p16+ or HPV16+	LSIL+ or p16+ or HPV16/33+
PPV	7.53 (5.06, 10.69)	8.44 (5.59, 12.12)	7.05 (4.74, 10.03)	7.69 (5.45, 10.49)	6.03 (4.32, 8.15)	6.21 (4.36, 8.53)	6.20 (4.38, 8.48)	5.88 (4.18, 8.02)	5.12 (3.67, 6.94)	5.19 (3.73, 7.00)
1-NPV	1.67 (0.87, 2.90)	1.79 (0.98, 2.98)	1.73 (0.90, 3.00)	0.64 (0.18, 1.64)	0.23 (0.01, 1.25)	0.95 (0.31, 2.20)	0.78 (0.21, 2.00)	0.65 (0.13, 1.89)	0.30 (0.01, 1.68)	0.00 (0.00, 1.15)

(c) Combinations of ASC-US+ cytology, p16 immunostaining and HPV genotypes for HPV-positive women (n = 1,091)						
	ASC-US+	ASC-US+ or p16+	ASC-US+ or HPV16+	ASC-US+ or HPV16/33+	ASC-US+ or p16+ or HPV16+	ASC-US+ or p16+ or HPV16/33+
N positive (%)	542 (49.68)	735 (67.37)	681 (62.42)	694 (63.61)	827 (75.80)	836 (76.63)
				CIN2+ (n = 92)		
N	80	91	85	86	91	92
Sensitivity	86.96 (78.32, 93.07)	98.91 (94.09, 99.97)	92.39 (84.95, 96.89)	93.48 (86.34, 97.57)	98.91 (94.09, 99.97)	100.00 (96.07, 100.00)
Specificity	53.75 (50.60, 56.88)	35.54 (32.56, 38.59)	40.34 (37.28, 43.46)	39.14 (36.10, 42.24)	26.33 (23.62, 29.17)	25.53 (22.85, 28.35)
PPV	14.76 (11.88, 18.03)	12.38 (10.09, 14.98)	12.48 (10.09, 15.20)	12.39 (10.03, 15.08)	11.00 (8.95, 13.34)	11.00 (8.96, 13.32)
1-NPV	2.19 (1.13, 3.79)	0.28 (0.01, 1.56)	1.71 (0.69, 3.49)	1.51 (0.56, 3.26)	0.38 (0.01, 2.09)	0.00 (0.00, 1.44)
				CIN3+ (n = 40)		
N	33	39	37	38	39	40
Sensitivity	82.50 (67.22, 92.66)	97.50 (86.84, 99.94)	92.50 (79.61, 98.43)	95.00 (83.08, 99.39)	97.50 (86.84, 99.94)	100.00 (91.19, 100.00)
PPV	6.09 (4.23, 8.44)	5.31 (3.80, 7.18)	5.43 (3.85, 7.41)	5.48 (3.90, 7.44)	4.72 (3.37, 6.39)	4.78 (3.44, 6.46)
1-NPV	1.28 (0.51, 2.61)	0.28 (0.01, 1.56)	0.73 (0.15, 2.12)	0.50 (0.06, 1.81)	0.38 (0.01, 2.09)	0.00 (0.00, 1.44)

[1]Includes lesions detected within 12 months of the first referral to colposcopy.

Abbreviations: 1-NPV, 1-negative predictive value; ASC-US+, positive atypical squamous cells of undetermined significance; CIN2+, cervical intraepithelial neoplasia grade 2+; CIN3+, cervical intraepithelial neoplasia grade 3+; HPV, human papillomavirus; HSIL+, positive high-grade squamous intraepithelial lesion; LSIL+, positive low-grade squamous intraepithelial lesion; PPV, positive predictive value.

Individual genotyping for the 13 hrHPV types positive by HC2 using GP5+/6+ primers were also examined. In order to limit overfitting issues, we defined a priori a selection procedure where markers had to separately have a positive predictive value (PPV) of at least 10% for CIN2+ to be considered in combination.

Sensitivity, specificity, PPV and false-negative reports (1-negative predictive value [1-NPV]) for CIN2+ and CIN3+ were computed for combinations of markers. CIN1 and normal were grouped as <CIN2 for all analyses. For individual genotyping a hierarchical ordering of HPV types was developed based on sequentially maximizing the PPV for the next HPV type after excluding women with multiple infections with types higher in the hierarchy.

Data availability

According to Italian law, anonymized data can only be made publicly available if there is no potential for the reidentification of individuals (https://www.garanteprivacy.it). The data underlying our study or ad hoc analyses are available on request to researchers who meet the criteria for access to confidential data. Requests should be addressed to the corresponding author.

Results

During Phase 2 of NTCC, 1,936 (7.9%) women from the experimental arm were positive for hrHPV by HC2 and 1,813 (93.6%) of these had colposcopy. Details of their follow-up are shown in Figure 1. Within this group, 1,547 women came from centers where samples were collected for cytology, p16 and genotyping. Biopsies were taken in 41.8% (647/1547) of women who had colposcopy and the remainder were considered negative. Complete test results for cytology, p16 and genotyping were available for 1,091 hrHPV-positive women (Fig. 1). Overall, 138 of these women were diagnosed with CIN2+ at baseline or during follow-up, including 60 with CIN3+ and 4 with invasive cervical cancer.

Disease detected at baseline (cross-sectional analysis)

At baseline, 92 women were diagnosed with CIN2+ and 40 of these had CIN3+. The sensitivity of HSIL+ cytology was only 53.3% (42.6, 63.7) for CIN2+ and 60.0% (43.3, 75.1) for CIN3+, but specificity for <CIN2 was very high (96.9% [95.6, 97.9]) (Table 1*a*). For LSIL+ cytology the sensitivity for CIN2+ increased to 73.9% (63.7, 82.5), but specificity was reduced to 69.6% (66.6, 72.4) (Table 1*b*), while for ASC-US+ cytology sensitivity increased further to 87.0% (78.3, 93.1) but specificity was reduced to 53.8% (50.6, 56.9) (Table 1*c*).

A total of 468 (42.9%) women were positive for p16 immunostaining. The sensitivity was 84.8% (75.8, 91.4) for CIN2+ and 90.0% (76.3, 97.2) for CIN3+, with a specificity for <CIN2 of 61.0% (57.9, 64.0) (Table 1*b*). p16 detected more CIN2+ than LSIL+ (73.9%), HPV16 (54.4%) or HPV16/33 (62.0%) (Table 1*b*).

HPV16 was the commonest genotype detected overall (N = 308, 28.2%) and after omitting the 11 multiple infections with types higher in the hierarchy (i.e., only HPV33, Supporting Information Table S1*a*), it had a similar PPV for CIN2+ as p16 (16.5 *vs.* 16.7%), but slightly lower than LSIL+ cytology (18.3%) (Table 1*b*). Ordering genotypes (with PPV > 10%) by decreasing PPV, HPV33 was ranked first with a PPV of 20.9% (10.0, 36.0) for CIN2+ (11.6% (3.9, 25.1) for CIN3+), followed by HPV16, 35, 59, 31 and 52 (Supporting Information Table S1*a*). However, the number of CIN2+ cases for HPV types 35, 59, 31 and 52 was very small. HPV18 had a lower discriminatory value with a PPV of only 7.3% (2.4, 16.1).

Combinations of cytology (ASC-US+, LSIL+, HSIL+), p16 positivity and genotyping for HPV types 16 and 33 were considered, as each had univariate PPV values for CIN2+ greater than 10% (Table 1, Fig. 2). Almost all CIN2+ disease was either LSIL+ or p16 positive (sensitivity 97.8% [92.4, 99.7] for CIN2+ and 97.5% [86.8, 99.9] for CIN3+), and only referring women positive for at least one of these makers reduced referrals by 40.7%, compared to referring all hrHPV-positive women (Table 1*b*). Very similar performance was seen if only p16 or HPV16-positive women were referred (Table 1*b*). Lower sensitivity was seen if only those who were HPV16+ or LSIL+ were referred (sensitivity 84.8% [75.8, 91.4] for CIN2+ and 87.5% [73.2, 95.8] for CIN3+), and in this case 51.7% of women would have been referred to colposcopy. Including ASC-US+ cytology increased sensitivity (ASC-US+ and/or HPV16+: 92.4% [85.0, 96.9] for CIN2+ and 92.5% [79.6, 98.4] for CIN3+), but at the expense of referring another 10.7% of the population (Table 1*c*). Adding HPV33 positivity only slightly improved performance compared to HPV16 alone, or with p16 or LSIL+ cytology (Table 1), but the number of cases was too small to draw conclusions. Figure 2 shows graphically the cross-sectional-added sensitivity and reduced specificity for combinations of p16 IHC, genotyping and LSIL+ cytology.

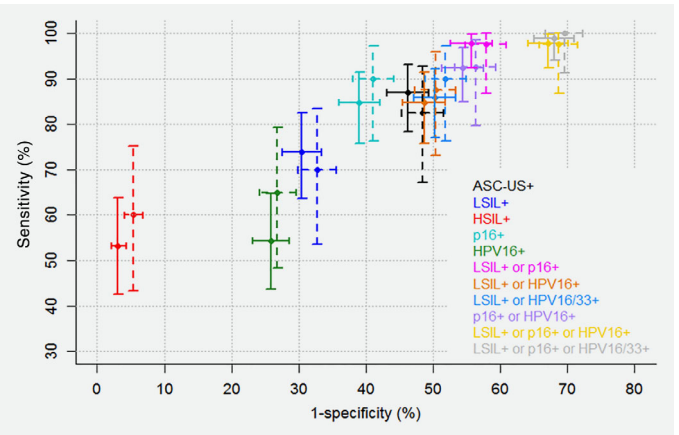

Figure 2. Sensitivity and specificity for combinations of triage tests for HPV-positive women. Solid line for CIN2+, dashed line for CIN3+. Only lesions detected at baseline included.

Table 2. Accuracy for detection of CIN2+ and CIN3+ by combinations of LSIL+ cytology, p16 immunostaining and HPV genotypes for HPV-positive women

	LSIL+	HPV16+	HPV16/33+	p16+	LSIL+ or p16+	LSIL+ or HPV16+	LSIL+ or HPV16/33+	p16+ or HPV16+	LSIL+ or p16+ or HPV16+	LSIL+ or p16+ or HPV16/33+
N positive (%)	372 (34.10%)	308 (28.23%)	397 (36.39%)	468 (42.90%)	647 (59.30%)	564 (51.70%)	581 (53.25%)	629 (57.65%)	761 (69.75%)	771 (70.67%)
CIN2+ (*n* = 138)										
N	90	69	78	108	125	112	114	123	131	132
Sensitivity	65.22 (56.65, 73.12)	50.00 (41.38, 58.62)	56.52 (47.82, 64.93)	78.26 (70.44, 84.83)	90.58 (84.43, 94.89)	81.16 (73.63, 87.31)	82.61 (75.24, 88.53)	89.13 (82.71, 93.79)	94.93 (89.83, 97.94)	95.65 (90.78, 98.39)
Specificity	70.41 (67.40, 73.29)	74.92 (72.04, 77.65)	66.53 (63.43, 69.52)	62.22 (59.06, 65.31)	45.23 (42.03, 48.45)	52.57 (49.34, 55.78)	51.00 (47.77, 54.22)	46.90 (43.70, 50.13)	33.89 (30.89, 37.00)	32.95 (29.97, 36.03)
PPV	24.19 (19.93, 28.88)	22.40 (17.87, 27.48)	19.65 (15.85, 23.90)	23.08 (19.33, 27.16)	19.32 (16.35, 22.58)	19.86 (16.64, 23.39)	19.62 (16.47, 23.09)	19.55 (16.52, 22.87)	17.21 (14.60, 20.09)	17.12 (14.53, 19.97)
1-NPV	6.68 (4.96, 8.75)	8.81 (6.92, 11.02)	8.65 (6.66, 10.99)	4.82 (3.27, 6.80)	2.93 (1.57, 4.95)	4.93 (3.25, 7.15)	4.71 (3.04, 6.92)	3.25 (1.83, 5.30)	2.12 (0.86, 4.32)	1.88 (0.69, 4.04)
CIN3+ (*n* = 60)										
N	37	36	39	51	56	50	51	54	57	58
Sensitivity	61.67 (48.21, 73.93)	60.00 (46.54, 72.44)	65.00 (51.60, 76.87)	85.00 (73.43, 92.90)	93.33 (83.80, 98.15)	83.33 (71.48, 91.71)	85.00 (73.43, 92.90)	90.00 (79.49, 96.24)	95.00 (86.08, 98.96)	96.67 (88.47, 99.59)
PPV	9.95 (7.10, 13.45)	11.69 (8.32, 15.81)	9.82 (7.08, 13.18)	10.90 (8.22, 14.08)	8.66 (6.60, 11.09)	8.87 (6.65, 11.52)	8.78 (6.61, 11.38)	8.59 (6.52, 11.05)	7.49 (5.72, 9.60)	7.52 (5.76, 9.62)
1-NPV	3.20 (2.04, 4.76)	3.07 (1.97, 4.53)	3.03 (1.88, 4.59)	1.44 (0.66, 2.72)	0.90 (0.25, 2.29)	1.90 (0.91, 3.46)	1.76 (0.81, 3.32)	1.30 (0.48, 2.81)	0.91 (0.19, 2.63)	0.63 (0.08, 2.24)

Includes all disease detected at baseline and follow-up (n = 1,091).

Abbreviations: 1-NPV, 1-negative predictive value; CIN2+, cervical intraepithelial neoplasia grade 2+; CIN3+, cervical intraepithelial neoplasia grade 3+; HPV, human papillomavirus; LSIL+, positive low-grade squamous intraepithelial lesion; PPV, positive predictive value.

As only one positive triage test is needed for referral, we looked at the order in which they might be applied to avoid unnecessary tests and reduce costs. As discussed below, this will depend on what is routinely available, but ignoring the routine availability of any results and costing details (which will depend on local policy), we considered an ordering based on reducing the overall number of tests performed. As p16 was most often positive (n = 468), the number of tests is minimized by doing this first, followed by either cytology (LSIL+) (179 additionally positive) or genotyping (HPV16 positive) (161 additionally positive), where the choice between them was marginal. With this approach a second (reflex) test would have been needed in only 57% of HPV-positive women. There was little gain in disease detection in doing all three tests.

Disease detected at follow-up (longitudinal)

A further 46 CIN2+ including 20 CIN3+ cases were diagnosed during follow-up (Supporting Information Table S1*b*). Of the additional CIN2+ cases, 6 (13.0%) were HSIL+, 22 (47.8%) LSIL+ and 35 (76.1%) ASC-US+, 30 (65.2%) were p16 positive and 19 (41.3%) HPV16 positive. One case of invasive cervical cancer was detected at 13 months during follow-up and this case had ASC-US cytology and was also positive for HPV16 and p16. She had no biopsy at the first colposcopy and did not receive 1-year follow-up screening in the organized program.

Including disease detected either at baseline or follow-up, only 13/138 cases (9.4%) of CIN2+ and 4/60 (6.7%) cases of CIN3+ were <LSIL and negative for p16 at entry (Table 2). Only six CIN2+ and one CIN3+ of these cases were positive for HPV16. If using p16 or HPV16 positivity for triage, 15 (10.9%) CIN2+ and 6 (10.0%) CIN3+ would have been false negative (Table 2). The 3-year sensitivity for different triage strategies is illustrated in Supporting Information Figure S1.

Including baseline and follow-up, the PPV for CIN2+ increased to 23.1% (19.3, 27.2) for p16 positivity (10.9% (8.2, 14.1) for CIN3+), and to 19.3% (16.4, 22.6) for LSIL+ or p16 positivity (8.7% (6.6, 11.1) for CIN3+). For p16 and HPV16 positivity, PPVs increased to 19.6% (16.5, 22.9) for CIN2+ and to 8.6% (6.5, 11.1) for CIN3+ (Table 2). Supporting Information Tables S2*a* and S2*b* show the diagnostic accuracy, including baseline and follow-up, for ASC-US+ and HSIL+ cytology in combination with p16 and genotyping.

Discussion

Testing for hrHPV is well known to be the most sensitive method currently available for primary screening. However, good triage tests are needed to improve specificity. There are several possibilities including cytology, HPV genotyping, p16 immunostaining and DNA methylation, but currently there is no consensus as to how best use them.

Peeters *et al.*[14] review data on the use of p16 with or without Ki-67 as triage for abnormal cytology and Wright *et al.*[15] compare p16/Ki-67 to cytology in HPV positive women in the Addressing the Need for Advanced HPV Diagnostics (ATHENA) trial. In both cases, support for p16 immunostaining is provided, but these studies only provided cross-sectional evaluation, without adequate follow-up. Our cross-sectional results are similar to those from the Primary ASC-US LSIL Marker Study (PALMS)[16] and the Papillomavirus Dumfries and Galloway (PaVDaG) study,[17] which showed much greater cross-sectional sensitivity for dual stained p16/Ki-67 than pap cytology or HPV16/18 genotyping, but again longitudinal follow-up and genotyping beyond Types 16 and 18 was not performed in either of these studies. Clarke *et al.*[18] also showed much better detection rates over a 5-year follow-up with dual p16/Ki-67 staining than for cytology. Wright *et al.*[15] have also shown increased sensitivity for detection of CIN2+ and CIN3+ for HPV-positive women when using p16/Ki-67 and/or HPV16/18 genotyping as a triage strategy compared to only cytology, but again longitudinal follow-up was neither reported nor was complete HPV genotyping performed.

A very recent cohort study[19] of women screened by cotesting with HPV and cytology also performed a supplementary research dual stain (p16/Ki67) test which was not used for management and included a 3-year follow-up. They considered the accuracy of cytology at an ASC-US cutoff and of genotyping for HPV16/18 (but not full genotyping). In this cohort, HPV-positive/cytology-negative women had no colposcopy if they were HPV negative at retesting, leaving the possibility of some verification bias. However, the sensitivity of dual staining was very similar to our findings suggesting that bias was minimal. The authors of our study concluded that extending screening intervals to 3 years in HPV16/18-negative women who are dual-stain negative was safe, but did not consider combinations of dual staining with cytology.

Our results indicate that HPV-positive women who are negative for both p16 immunostaining and LSIL+ cytology have a very low risk of CIN2+ (0.5% [0.1, 1.6]) and especially CIN3+ (0.2% [0.01, 1.3]) at an initial colposcopy, and this risk remains low after 3-year follow-up (2.9% [1.6, 5.0] and 0.9% [0.3, 2.3], respectively). No cancer was diagnosed during the 3-year follow-up in women negative for either of these triage combinations.

The high sensitivity of the triage strategies proposed would allow retesting triage-negative women after 3 years, so avoiding a substantial number of short-term repeat tests, and reducing costs, anxiety for women, and loss to follow-up. The triage strategies proposed would also avoid a substantial number of colposcopies. Immediate referral to colposcopy appears to be slightly higher (59.3% of HPV positives when referring either p16 positive or LSIL+ women, 57.7% when referring HPV-positive women for either p16 or HPV16) compared to the currently widely used ASC-US+ cytology only (49.7%). However, it is lower than the current U.S. recommendation[20,21] of referral

of either HPV16+ or ASC-US+ women, which, in our study, would have led to an even higher 62.4% immediate referral without any gain in sensitivity.

However, only 10–15% of HPV infections persist for at least 3 years *versus* about 40% positivity after 1 year.[22] Therefore, retesting triage negative women for HPV after 3 years and referring to colposcopy those remain HPV positive will strongly reduce the number of delayed colposcopies compared to referral for those persistent after 1 year. We estimate our approach would lead to colposcopy referral in about 62–63% of HPV-positive women overall, *versus* 80% if all ASC-US+ or HPV16+ women were referred immediately, and the negatives were managed by annual repeat HPV testing. If just ASC-US+ women were referred immediately and negatives were retested after 1 year, the referral rate would be 70%.

It must, however, be kept in mind that short-term follow-up is needed for women who have a first negative colposcopy. The number of tests needed is expected to be proportional to the overall colposcopy referral rate, therefore, it will also be reduced with our proposed strategy.

Reducing the overall number of tests is important but this needs to be viewed in light of the routine use of cytology, especially when cotesting with both HPV and cytology is practiced. HPV16/18 genotyping is also automatically provided with some HPV tests. Genotyping also has the advantage of being objective and reproducible, which is not the case for low-grade cytology. Costs are country specific and a full economic evaluation is needed to define the most cost-effective strategy.

HPV16 is more common than the other oncogenic types, both in high-grade CIN and invasive cancers.[23] Currently, most commercially available HPV tests only genotype for HPV16/18. However, it has been shown that other hrHPV types, particularly HPV31 and 33 are associated with a high rate of high-grade lesions and PPVs leading to greater sensitivity.[8,23,24] HPV18 and 45 are not strongly associated with CIN3+ in the next 3 years, but are the second and third most common types in invasive cervical carcinoma, and are specifically associated with glandular intraepithelial lesions and adenocarcinoma. Thus, while it may be useful to test for HPV18 (and HPV45), management of positive women needs to be different, as disease yield within 3 years is low and only persistent infection after 3 years would seem to warrant colposcopy with a more careful exploration of the endocervical canal to look for lesions not apparent in routine colposcopy. A more conservative strategy would be to manage these women with short-term (e.g., 1 year) repeat testing, but this requires further evaluation in other cohorts.

The NTCC trial was a large, population-based study, nested in routine organized screening in a low-risk population. Over 70% of eligible women were enrolled in the study,[1] suggesting that results are applicable to routine practice. One of the main strengths of our study was the referral to colposcopy (with high participation) of all HPV-positive women, and the 3-year follow-up of women who did not have CIN2+ initially detected at colposcopy. Such a design minimizes the risk of verification bias. In almost all cases the histopathologic endpoint was determined by a central review blinded to the HPV test, cytology, p16 and genotyping results. In addition, we also searched cancer registries and pathology units for lesions detected outside participating programs.

We obtained specimens for genotyping and cytology results from a large proportion of the HPV positive women in the centers included in the study. A proportion of p16 specimens from women without biopsy at baseline were randomly discarded to reduce costs, and the remaining missing samples were the result of delayed start in sample collection and can reasonably be considered as "missing at random." Samples for cytology and p16 testing were read in the knowledge of HPV positivity. This is known to increase both true-positive and false-positive rates[7,25] and can be seen in the lower specificity observed here compared to other studies,[15] but reflects the proposed use and is thus an advantage of our study design. A sample of cytology slides (n = 852) were also reviewed externally, blind to histology,[7] and results for sensitivity were very similar, but specificity was higher than for the original interpretation (data not shown). Genotyping and p16 were blind to histology and performed in different laboratories.

Dual staining for p16 and Ki-67 is now widely used in order to minimize subjectivity in interpretation, but was not available when we performed immunostaining. Results suggest similar sensitivity as for stand-alone p16, but better specificity.[7] However, the sensitivity of p16 could have been underestimated due to the use of cytospin preparations instead of full-size standard LBC slides. We found that extended genotyping to detect HPV33 improved performance, but the gain was small.

In conclusion, our data suggest that p16 immunostaining combined with either cytology or some level of genotyping should be used to triage HPV-positive women. This can maintain high sensitivity and lead to a substantial reduction in the number of women referred for colposcopy or managed by short-term repeat testing.

Acknowledgements

Our study was supported by the Italian Ministry of Health (Progetto Speciale "Valutazione di nuove tecnologie per lo screening del cervicocarcinoma—follow-up" and l 138-2004 progetto "NTCC—nuove tecnologie per lo screening del cervicocarcinoma"), by the European Union (COHEAHR project, FP7 grant agreement no. Health-F3-2013-603019) and by the Associazione Italiana per la Ricerca sul Cancro (AIRC, project IG14119). Sponsors 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 the study.

Conflict of interest

P.G.R. as former PI of an independent study funded by the Italian ministry of health conducted negotiations with Roche, Hologic, Becton Dickinson to obtain reagents at reduced price or for free. J.C. receives grants for his institute from Qiagen, Becton Dickinson, Genera Biosystems, Hologic, Genefirst and TrovaGene. F.C. is part of the advisory board of Becton Dickinson. A.G.T. and L.D.M. as molecular lab coordinators, in an independent study funded by the Italian Ministry of Health, are conducting negotiations with Becton Dickinson to obtain reagents for free and a scholarship holder to perform virus genotyping. All other authors reported no conflict of interest.

Disclaimer

The authors alone are responsible for the views expressed in this paper and they do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated.

References

1. 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.
2. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet* 2014;383: 524–32.
3. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine* 2012;30:88–99.
4. Waller J, McCaffery K, Kitchener H, et al. Women's experiences of repeated HPV testing in the context of cervical cancer screening: a qualitative study. *Psychooncology* 2007;16:196–204.
5. Ronco G, Franceschi S, Segnan N. HPV16 and HPV18 genotyping in cervical cancer screening. *Lancet Oncol* 2011;12:831–2.
6. 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.
7. Bergeron C, Giorgi-Rossi P, Cas F, et al. Informed cytology for triaging HPV-positive women: substudy nested in the NTCC randomized controlled trial. *J Natl Cancer Inst* 2015;107:dju423.
8. Del Mistro A, Adcock R, Carozzi F, et al. Human papilloma virus genotyping for the cross-sectional and longitudinal probability of developing cervical intraepithelial neoplasia grade 2 or more. *Int J Cancer* 2018;143:333–42.
9. McCredie MR, Sharples KJ, Paul C, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol* 2008;9:425–34.
10. 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.
11. 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.
12. Ronco G, Giorgi-Rossi P, Carozzi F, et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst* 2008;100: 492–501.
13. van den Brule AJC, Pol R, Fransen-Daalmeijer N, et al. GP5+/6+ PCR followed by reverse line blot analysis enables rapid and high-throughput identification of human papillomavirus genotypes. *J Clin Microbiol* 2002;40:779–87.
14. Peeters E, Wentzensen N, Bergeron C, et al. Meta-analysis of the accuracy of p16 or p16/Ki-67 immunocytochemistry versus HPV testing for the detection of CIN2+/CIN3+ in triage of women with minor abnormal cytology. *Cancer Cytopathol* 2019;127:169–80.
15. Wright TC Jr, Behrens CM, Ranger-Moore J, et al. Triaging HPV-positive women with p16/Ki-67 dual-stained cytology: results from a substudy nested into the ATHENA trial. *Gynecol Oncol* 2017;144:51–6.
16. Ikenberg H, Bergeron C, Schmidt D, et al. Screening for cervical cancer precursors with p16/Ki-67 dual-stained cytology: results of the PALMS study. *J Natl Cancer Inst* 2013;105:1550–7.
17. Stanczuk GA, Baxter GJ, Currie H, et al. Defining optimal triage strategies for hrHPV screen-positive women-an evaluation of HPV 16/18 genotyping, cytology, and p16/Ki-67 cytoimmunochemistry. *Cancer Epidemiol Biomarkers Prev* 2017;26:1629–35.
18. Clarke MA, Cheung LC, Castle PE, et al. Five-year risk of cervical precancer following p16/Ki-67 dual-stain triage of HPV-positive women. *JAMA Oncol* 2019;5:181–6.
19. Wentzensen N, Clarke MA, Bremer R, et al. Clinical evaluation of human papillomavirus screening with p16/Ki-67 dual stain triage in a large organized cervical cancer screening program. *JAMA Intern Med* 2019;179:881–8.
20. Saslow D, Solomon D, Lawson HW, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening guidelines for the prevention and early detection of cervical cancer. *Am J Clin Pathol* 2012;137:516–42.
21. Huh WK, Ault KA, Chelmow D, et al. Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance. *Obstet Gynecol* 2015;125:330–7.
22. Rositch AF, Koshiol J, Hudgens MG, et al. Patterns of persistent genital human papillomavirus infection among women worldwide: a literature review and meta-analysis. *Int J Cancer* 2013; 133:1271–85.
23. Cuzick J, Wheeler C. Need for expanded HPV genotyping for cervical screening. *Papillomavirus Res* 2016;2:112–5.
24. Cuzick J, Ho L, Terry G, et al. Individual detection of 14 high risk human papilloma virus genotypes by the PapType test for the prediction of high grade cervical lesions. *J Clin Virol* 2014; 60:44–9.
25. Wright TC Jr, Stoler MH, Aslam S, et al. Knowledge of patients' human papillomavirus status at the time of cytologic review significantly affects the performance of cervical cytology in the ATHENA study. *Am J Clin Pathol* 2016;146: 391–8.