



A cohort study of cervical screening using partial HPV typing and cytology triage

Mark Schiffman¹, Noorie Hyun¹, Tina R. Raine-Bennett², Hormuzd Katki¹, Barbara Fetterman³, Julia C. Gage¹, Li C. Cheung⁴, Brian Befano⁴, Nancy Poitras³, Thomas Lorey³, Philip E. Castle⁵ and Nicolas Wentzensen¹

¹ National Cancer Institute Division of Cancer Epidemiology and Genetics, Rockville, MD

² Kaiser Permanente Northern California Division of Research, Oakland, CA

³ Kaiser Permanente Northern California Regional Laboratory, Berkeley, CA

⁴ Information Management Services, Calverton, MD

⁵ Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY

HPV testing is more sensitive than cytology for cervical screening. However, to incorporate HPV tests into screening, risk-stratification (“triage”) of HPV-positive women is needed to avoid excessive colposcopy and overtreatment. We prospectively evaluated combinations of partial HPV typing (Onclarity, BD) and cytology triage, and explored whether management could be simplified, based on grouping combinations yielding similar 3-year or 18-month CIN3+ risks. We typed ~9,000 archived specimens, taken at enrollment (2007–2011) into the NCI-Kaiser Permanente Northern California (KPNC) HPV Persistence and Progression (PaP) cohort. Stratified sampling, with reweighting in the statistical analysis, permitted risk estimation of HPV/cytology combinations for the 700,000+ woman KPNC screening population. Based on 3-year CIN3+ risks, Onclarity results could be combined into five groups (HPV16, else HPV18/45, else HPV31/33/58/52, else HPV51/35/39/68/56/66/68, else HPV negative); cytology results fell into three risk groups (“high-grade,” ASC-US/LSIL, NILM). For the resultant 15 HPV group-cytology combinations, 3-year CIN3+ risks ranged 1,000-fold from 60.6% to 0.06%. To guide management, we compared the risks to established “benchmark” risk/management thresholds in this same population (*e.g.*, LSIL predicted 3-year CIN3+ risk of 5.8% in the screening population, providing the benchmark for colposcopic referral). By benchmarking to 3-year risk thresholds (supplemented by 18-month estimates), the widely varying risk strata could be condensed into four action bands (very high risk of CIN3+ mandating consideration of cone biopsy if colposcopy did not find precancer; moderate risk justifying colposcopy; low risk managed by intensified follow-up to permit HPV “clearance”; and very low risk permitting routine screening.) Overall, the results support primary HPV testing, with management of HPV-positive women using partial HPV typing and cytology.

Key words: cervix, HPV, screening, cancer, cytology

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

Disclosure Statement: Drs. Schiffman and Gage report that NCI has received testing of specimens from the PaP cohort at reduced or no cost, for NCI-led studies from Roche and BD. Dr. Wentzensen reports receiving testing of specimens at reduced or no cost for NCI-led studies from Roche, BD, and Hologic. Dr. Castle has received commercial HPV tests for research at a reduced or no cost from Roche, Qiagen, Norchip, Arbor Vita Corporation, BD and mtm. He has been compensated financially as a member of a Merck Data and Safety Monitoring Board for HPV vaccines. Dr. Castle has been paid as consultant for Gen-Probe/Hologic, Roche, Cepheid, ClearPath, Guided Therapeutics, Teva Pharmaceuticals, Gentcel, Inovio and GE Healthcare. Dr. Castle has received honoraria as a speaker for Roche and Cepheid.

DOI: 10.1002/ijc.30375

History: Received 21 Apr 2016; Accepted 15 July 2016; Online 10 Aug 2016

Correspondence to: Mark Schiffman, Room 6E544, 9609 Medical Center Drive, Rockville, MD 20850, USA, Tel.: 240-276-7259, E-mail: schiffmm@exchange.nih.gov

Human papillomavirus (HPV) infection with approximately a dozen high-risk types is the necessary cause of virtually all cases of cervical cancer and precancer.¹ Precancer (diagnosed histologically as CIN3/AIS) is the optimal target of cervical screening, because it can be diagnosed and treated to prevent cancer morbidity and mortality.² HPV testing is more sensitive for the detection of cervical precancer than cytology and therefore permits longer screening intervals.^{3,4}

However, specificity remains an issue; high-risk HPV infections are quite commonly acquired by sexual activity and, although carcinogenic potential varies by type, even the highest-risk types of HPV usually are controlled immunologically or otherwise spontaneously “clear” within 1–2 years.² To avoid possibly harmful overtreatment of many women needing no intervention, those testing positive for HPV must be further stratified (“triaged”) to identify likely prevalent or incipient cervical precancer.⁵ Therefore, while it is increasingly accepted that cervical screening should incorporate HPV testing rather than cytology alone, there is still substantial disagreement as to the details, specifically, over how best to benefit from the sensitivity of HPV testing while minimizing

What's new?

Cervical cancer screening includes two elements: cytology and HPV testing. But many women who test positive for HPV won't need further examination; most HPV infections don't lead to cancer. How best to identify those at risk? These authors created a system to simplify risk assessment. By conducting a cohort study to document the likelihood of patients developing CIN3 within three years, they evaluated different patterns of screening results. They were able to stratify risk profiles generated by HPV typing and cytology into four "action bands," or further screening recommendations, thus balancing the sensitivity of HPV typing with simplicity of implementation.

overtreatment.^{4,6} In many countries, primary HPV testing alone is being considered, possibly after a period of cotesting with HPV and cytology together.⁷ In contrast, U.S. guidelines tend to recommend continued cotesting with HPV and cytology together, although cytology remains an acceptable option³ and HPV testing alone is now FDA-approved for one assay.⁸

When screening involves HPV testing in the form of cotesting or HPV testing alone, an important secondary question involves the usefulness of HPV typing, that is, whether it is worth distinguishing between the individual HPV types in the high-risk group to help guide management of women screening HPV-positive.^{9–11} In the U.S., FDA-approved HPV test systems vary in the degree of typing provided: HC2 (Qiagen, Germantown, MD) provides a pooled result and no typing; Cervista (Hologic, Marlborough, MA) and cobas (Roche Molecular Systems, Pleasanton, CA) identify HPV16 and HPV18 (the most important types that account for more than half of cervical cancer cases), and Aptima (Hologic, Marlborough, MA) separately identifies RNA of HPV16 and pools HPV18 with a closely related type, HPV45.

A fifth HPV DNA test system that has recently completed the enrollment phase of an FDA-registration trial is the BD Onclarity™ HPV Assay.¹² Onclarity has nine typing channels: HPV16, HPV18, HPV31, HPV45, HPV51, HPV52, HPV33/58, HPV35/39/68 and HPV56/59/66.^{12,13} It remains unclear how much of this individual and small-group partial typing is useful for guiding clinical management of HPV-positive women.¹⁰ A second question is how cytology can be combined usefully with HPV typing to estimate jointly the risk of precancer/cancer.⁸ A third question is, given the resultant detailed risk information, how we can condense the detail to create simple and practical management strategies.¹⁴ We addressed these questions as part of an ongoing observational cohort study of >50,000 women, called the HPV Persistence and Progression (PaP) cohort, underway since the beginning of 2007 at Kaiser Permanente Northern California (KPNC) in collaboration with the National Cancer Institute (NCI).¹⁵

Material and Methods**Study population**

Cytology and HPV (HC2) cotesting at 3-year intervals was introduced at KPNC in 2003.¹⁶ Since then, more than 1.5

million women have been cotested, making the KPNC experience one of the largest applications of HPV testing in cervical screening. To establish the PaP cohort (enrollment from January 2007–January 2011), residual HPV test specimens (Specimen Transport Medium, Qiagen) remaining after HC2 testing were retrieved, neutralized and frozen from ~55,000 women receiving both cytology and HPV testing.^{17,18} The enrollment cotest was defined as the first one with a stored residual STM specimen during the collection period.

The goal of the PaP cohort is primarily to study the molecular epidemiologic predictors of precancer among high-risk HPV-infected women.⁹ Approximately 45,000 of the total enrollment were residual aliquots of HC2-positive specimens, representing >3/4 of all specimens testing HC2-positive at KPNC during the enrollment interval (of almost a million women cotested during that interval). Residual specimens from a random group of ~10,000 HC2-negative women seen during the same period were collected for comparison. Opt-out consent for use of residual specimens and related health information was obtained; 8% of women opted out of the study. Personal identifiers were removed from specimens; study IDs that could be linked anonymously to clinical data were assigned. Clinical outcome data, including subsequent screening and follow-up cervical cytologic and HPV test results, as well as cervical histologic outcomes, were obtained from electronic health records. The cohort study has been re-approved by both KPNC and NCI IRB review committees yearly.

Stratified sampling of specimens for onclarity partial typing

The present analysis considers cotest results for PaP cohort enrollment specimens from women aged 30–65, the cohort was screened according to KPNC guidelines at 3-year intervals. At the time of analysis, follow-up data on histologic outcomes were available through 2014. The study population from which this analysis was drawn is shown in Supporting Information Figure 1. To evaluate Onclarity partial typing, we restricted to those obtaining routine screening, as opposed to women being tested for known prior cervical screening abnormalities and/or post treatment. Of the targeted population of >740,000 women aged 30–65 with first (enrollment) cotest results from January 2007 to January 2011, approximately 5% of the eligible women were HC2-positive at enrollment. We successfully stored 82% of those testing positive

along with a randomly selected 0.7% of specimens testing HC2-negative. To study the 3-year CIN3+ risks for HPV typing at BD, we conducted a second phase of stratified random sampling, included 9,258 specimens (Supporting Information Table 1). The focus was on HC2-positive women. We restricted retesting of HC2-negative specimens to a random group of 500, then supplemented the random group with all 94 available HC2-negative enrollment specimens in the PaP cohort, linked in the follow-up to rare outcomes suggesting elevated cancer risk, specifically, CIN2+ histology or “high grade” cytology (the latter including ASC-H, AGC, HSIL, AIS, or cancer results).

In Supporting Information Table 1 and its accompanying text, we detail the sampling fractions and weights used to reconstitute the larger KPNC screening population from the stratified random sample selected for Onclarity testing. Because precancer outcomes proved so rare among HC2-negative women, the supplemented 594-woman sample could not provide stable estimates of the very low risk among HPV-test negative populations. To obtain stable risk estimates, when risk was so low, we made use of the close agreement between HC2 and Onclarity regarding HPV negativity.¹² We estimated the low CIN3+ risks among HPV-negative women by using the full data from all 685,577 HC2-negative women at KPNC enrolled during 2007–2013.

BD Onclarity testing

The BD OnclarityTM HPV Assay is a multiplex real-time PCR assay that targets HPV type-specific E6 or E7 sequences.¹³ An FDA trial for U.S. licensure is currently underway; the test is already available in several other countries with full reporting of all nine typing channels and DNA adequacy results.¹⁹

Masked testing with Onclarity was performed at BD (Sparks, MD). Previously, 25 μ L of the STM specimen was found to give a similar Ct response to 0.5 mL of liquid-based cytology medium and had a similar analytic sensitivity to Hybrid Capture 2.¹² The equivalent aliquot of the neutralized (and thus more dilute) STM specimen was used in this study. DNA extraction, PCR amplification and detection were performed on the BD Viper LT System, and the data derived from the nine typing channels were analyzed and interpreted at NCI.

Statistical analysis

Like all observational studies of cervical HPV infection drawn from routine clinical practice, the KPNC cohort has two prominent features: the time-to-worst histology outcome is interval-censored between irregular visits; and many of the women with CIN3+ have prevalent disease at enrollment that might or might not be diagnosable if colposcopy is performed. Ignoring interval censoring could lead to invalid inferences,^{20–22} and excluding the subjects with prevalent disease would yield incomplete risk estimates. We developed a prevalence-incidence model to simultaneously account for

events occurring prior to or at enrollment and those found as incident in the follow-up for women that returned. We considered prevalent disease as a left-censored point-mass at time zero because we were not considering the time in the past when the outcome of CIN3/AIS detected at enrollment might have arisen. Logistic and Cox modeling were employed for prevalence and incidence risk estimates, respectively, except as noted.

The sampling design was a two-phase stratified simple random sample design as detailed in Supporting Information Table 1; in the 1st phase, the five factors were used to construct informative strata among the population of all screened women, and in the 2nd phase, stored specimens representing each stratum were randomly sampled. In total, 8,664 HC2-positive specimens stored in the PaP collection were randomly selected for HPV typing. The weight to estimate risks for the PaP cohort was the inverse of the sampling fraction and we applied the weighted logistic-Cox model to the sampled data. To apply a normal asymptotic distribution for the test statistic induced by a logistic-Cox mixture model, we approximated the baseline hazard function in the Cox model by using cubic basis-splines. The cubic basis-spline is a piecewise polynomial function of degree ≤ 3 in a variable of time. The places where the piecewise functions meet are called knots. To approximate a hazard function, one needs to decide the number and locations of knots over time. The number of knots ranged from 7 to 9 and the placements were selected by the quantiles of uniquely ordered observation times. To justify good approximation of the baseline curve, we compared the cubic basis-spline-based curve with the non-parametric maximum likelihood estimates. We employed the Iterative-Convex-Minorant algorithm^{23,24} to obtain the absolute risk estimates from semiparametric and cubic basis-spline-based logistic-Cox models, respectively.

For the first major analysis, we conducted hierarchical analyses to group HPV genotypes into high-to-low order of 3-year cumulative CIN3+ risk (including baseline and follow-up risk), with the assumption of no effect modification between HPV genotypes. At each step of the hierarchical analysis, we excluded the group with the higher risk, as previously described,^{9,10} to determine the optimal order (and marginal utility) of introducing additional type channels for risk stratification. We did not correct the type 1 error for multiple comparisons because the hierarchical analyses were done for exploratory reasons.

In the second part of the analysis, we deduced the logical management of each combination of HPV risk group and cytology, as might be recommended by U.S. consensus guidelines. HPV genotypes were grouped into risk groups, which were combined in turn with grouped cytology results to form combined HPV-cytology risk strata. High-grade cytology included ASC-H and AGC as well as HSIL/cancer. The logical management of each risk stratum was deduced by comparison with established U.S. risk benchmarks. Benchmarks are the levels of risk that, by U.S. guidelines committee

Table 1. Hierarchical analysis by onclarity typing channel of 3-year cumulative risk of CIN3 or worse

HPV	Positive				Negative				Total
	No of obs.	3-year risk (%)	95% CI		No of obs.	3-year risk (%)	95% CI		
HPV16	1,564	21.9	20.1	23.7	6,769	4.2	3.9	4.5	8,333
else HPV18	494	11.5	9.2	13.8	6,275	3.7	3.4	3.9	6,769
else HPV33/58	631	8.6	6.9	10.3	5,644	3.3	3	3.5	6,275
else HPV31	766	8.1	6.6	9.5	4,878	2.5	2.3	2.8	5,644
else HPV45	324	5.4	3.8	7.0	4,554	2.1	1.9	2.3	4,878
else HPV52	823	5.6	4.4	6.7	3,731	1.8	1.5	2.0	4,554
else HPV51	536	2.9	1.9	3.9	3,195	1.7	1.4	1.9	3,731
else HPV39/68/35	1,201	2.0	1.5	2.5	1,994	1.4	1.1	1.7	3,195
else HPV59/56/66	1,047	1.5	1.0	1.9	947	1.4	1.0	1.8	1,994
HC2 negative	685,577	0.076	0.069	0.085					

consensus, dictate each recommended management choice, for example, colposcopy for LSIL or intensified retesting for ASC-US.^{25,26} By comparison to benchmarks, we condensed the management of the many risk strata into a simplified management strategy.¹⁴

In ancillary analyses, we used CIN2+ as a clinically relevant outcome measure, because CIN2+ is typically treated (although the diagnosis of CIN2 is not reproducible and many cases of CIN2 are likely to regress without treatment).²⁷ We also considered the invasive cancers separately, because of their rarity and importance.

Results

The median age of the 8,664 enrollment HPV-positive women tested by Onclarity after exclusions was 39.0 years (mean 41.0). The median follow-up time for controls after the enrollment test (in the KPNC data file used for this analysis, which extended through 2014) was 3.5 years (mean 3.2 years, IQR 1.2–4.9, maximum 7.9 years).

Table 1 shows the 3-year risks of CIN3+ by HPV status, estimated by type channel in a hierarchical manner. The hierarchical analysis of 3-year CIN3+ risks by Onclarity typing channels confirmed the uniquely high risk of HPV16, which was both common (approximately one-fifth of all positives) and associated with a 21.9% 3-year risk of CIN3+ when found at enrollment. In the absence of HPV16, HPV18 predicted the second greatest risk (11.5%). A group of type channels (HPV33/58, HPV31, HPV45 and HPV52) predicted similar, intermediate risk of CIN3+ (5.4–8.6%), while the remaining channels, including seven types that proved to be common in aggregate (HPV51, HPV35/39/68, HPV56/59/66), yielded lower risks (1.5–2.9%).

Of note, as shown in Table 1, a total of 947 HC2-positive women were negative when tested by Onclarity, and their 3-year risk of CIN3+ was 1.4%, not meaningfully different than the lowest-risk Onclarity-positive group (1.5%, *Z*-test = 0.47, *p*-values = 0.68). This group of women with

discrepant HC2/Onclarity would have been classified as negative in Onclarity primary screening.

The hierarchical analysis suggested that Onclarity-positive typing channels could be combined into four risk groups (for completeness, two other groups are also shown: HC2-positive/Onclarity-negative, and HC2 negative). We deviated from the otherwise strictly hierarchical empirical groupings based on CIN3+ risk, and grouped HPV45 with HPV18. We judged that this combination was logical based on the close phylogenetic similarity of HPV45 and HPV18,²⁸ the known importance of HPV45 worldwide in invasive cancers,²⁹ and the shared associations with glandular lesions, especially adenocarcinoma.^{29,30}

Table 2 addresses the use of cytology to stratify risk following HPV testing that yielded the type groups. Within each separate type group, the HC2-positive/Onclarity-positive group, and the HPV-negative group, we estimated the risks stratified by cytologic result grouped as NILM, ASC-US/LSIL and high-grade (HSIL, AIS, cancer, ASC-H and AGC). The decision to combine ASC-US and LSIL was supported by the lack of statistically significant differences in histologic outcomes when these two results were compared, for each of the six HPV test-defined strata (data not shown). Inclusion of AGC and ASC-H in the high-grade group was an imperfect solution to categorizing these rare, possibly worrisome results.³¹ Compared to HSIL+ alone, risks for all other high-grade results combined were lower.

As shown in Table 2, grouping of cytologic results into three risk groups created 18 HPV-cytology risk strata (only 12 of which would have been present if Onclarity primary screening had been used, as Onclarity-negative results would not trigger cytology). The combinations of HPV group type and cytology predicted widely varying 3-year risks ranging about 1,000-fold from 60.6% for HPV16 high-grade cytology to 0.056% for HPV-negative NILM. The great majority of women (95.0% in the overall 30–65 KPNC screening population) tested HPV-negative (by HC2). As mentioned above,

Table 2. 18-Month and 3-year cumulative risks of CIN3+ by HPV type groups and cytology in the NCI/KPNC PaP cohort

Onclarity result	Cytology	No. of obs.	% Of onclarity positives	18-Month risk (%)	95% Lower limit (%)	95% Upper limit (%)	3-Year risk (%)	95% Lower limit (%)	95% Upper limit (%)	Management
HPV 16	Overall	1,564	21.2%	20.0	19.4	20.7	21.9	20.1	23.7	
	High-grade	506	6.8%	59.1	58.6	59.6	60.6	59.5	61.6	Diagnostic cone?
	LSIL/ASCUS	574	7.8%	15.1	14.6	15.6	17.9	16.2	19.5	Colposcopy
	NILM	484	6.6%	10.7	9.1	12.4	13.8	11.9	15.6	Colposcopy
else HPV 18/45	Overall	843	11.4%	8.3	7.9	8.8	9.0	7.5	10.5	
	High-Grade	235	3.2%	40.6	40.1	41.1	40.9	39.7	42.1	Colposcopy
	LSIL/ASCUS	250	3.4%	6.0	5.7	6.4	7.1	5.6	8.6	Colposcopy
	NILM	358	4.8%	3.3	2.3	4.2	4.4	3.2	5.6	1 Year vs. Colpo
else HPV31/33/58/52	Overall	2,195	29.7%	6.0	5.9	6.1	7.3	6.5	8.1	
	High-grade	467	6.3%	34.5	34.1	34.8	35.0	34.2	35.7	Colposcopy
	LSIL/ASCUS	850	11.5%	5.4	5.1	5.8	5.7	5.0	6.4	Colposcopy
	NILM	878	11.9%	2.9	1.8	3.9	4.0	3.2	4.7	1 Year
else HPV51/35/39/68/56/59/66	Overall	2,784	37.7%	1.6	1.4	1.8	2.0	1.7	2.3	
	High-Grade	321	4.5%	13.6	13.0	14.1	13.6	11.8	15.3	Colposcopy
	LSIL/ASCUS	1,123	15.2%	1.5	1.4	1.6	2.0	1.7	2.4	1 Year ³
	NILM	1,340	18.1%	0.8	0.0	1.6	1.2	6.2	8.7	1 Year
else Onclarity neg but HC2 pos ¹	Overall	947	–	1.2	1.1	1.3	1.4	0.9	1.9	
	High-grade	112	–	14.4	11.8	17.1	14.4	12.2	16.7	Colposcopy ¹
	LSIL/ASCUS	420	–	1.2	1.1	1.3	1.5	1.0	2.0	1 Year ¹
	NILM	415	–	0	0	0.8	0.5	0.0	0.9	1 Year ¹

Table 2. 18-Month and 3-year cumulative risks of CIN3+ by HPV type groups and cytology in the NCI/KPNC PaP cohort (Continued)

Onclarity result	Cytology	No. of obs.	% Of onclarity positives	18-Month risk (%)	95% Lower limit (%)	95% Upper limit (%)	3-Year risk (%)	95% Lower limit (%)	95% Upper limit (%)	Management
HC2 negative	Overall	685,577	–	0.08	0.07	0.08	0.08	0.07	0.08	
	High-grade	99	–	35.9	26.1	48.0	49.6	40.4	59.8	Colposcopy ¹
	LSIL/ASCUS	15,616	–	0.33	0.24	0.44	0.33	0.24	0.44	3 Years
	NILM	667,942	–	0.06	0.05	0.06	0.06	0.05	0.06	5 Years ²

¹These would be categorized as HPV-negative if tested by Onclarity alone, and rescreened at 3 years. We had typing data from previous studies for all of the CIN3+ cases in this Onclarity-negative, HC2-positive discrepant group (31 CIN3, 1 AIS, 3 cancers). Most had been called HPV-positive by at least one assay in prior studies (21/31 CIN3, 1/1 AIS, 2/3 tested cancers), but about half of these were called positive for types not targeted by Onclarity (e.g., HPV53 for 1 cancer, and HPV34 or HPV73 for the other HPV-positive cancer). Of note, this study did not capture sufficiently the converse discrepancy of Onclarity-positive, HC2-negative test results.

²5 Years by guidelines for HPV-negative NILM, but would be rescreened at 3 years if HPV primary screening used.

³The lowest risk HPV-positive stratum, with cytology of ASC-US/LSIL, might switch from colposcopy to 1-year return due to risk lower than ASC-US cytology, HPV status unknown.

The benchmark risk thresholds for this population according to U.S. management guidelines, based on cytologic results alone, are: For 18-month risks: HSIL+ = 50.6% (consider diagnostic cone if colposcopy is not CIN2+); LSIL = 5.3% (colposcopy); ASC-US = 2.0% (1-year return testing); NILM = 0.13% (routine return screening at 3 years). For 3-year risks: HSIL+ = 52.0% (consider diagnostic cone if colposcopy is not CIN2+); LSIL = 5.8% (colposcopy); ASC-US = 2.3% (1-year return testing); NILM = 0.17% (routine return screening at 3 years).

Abbreviations: ASC-US: atypical squamous cells of undetermined significance; atypical glandular cells; adenocarcinoma *in situ*, or cancer cytology results; high-grade: high-grade intraepithelial lesion (HSIL) and also atypical squamous cells, rule out HSIL; LSIL: low-grade squamous intraepithelial lesion; NILM: negative for intraepithelial lesion or malignancy.

risk of CIN3+ was very low among all HPV-negative women, except among the extremely small percentage with high-grade cytology (99 (0.014%) of 685,577 HC2 negative), who had 3-year CIN3+ risk of 49.6%. However, among Onclarity-positive women, there was a substantial dispersion of test results and predicted risks across the HPV type-cytology categories, from 1.2% to 60.6%.

Although we emphasized 3-year risk estimates, the difficult decision as to whether to refer immediately to colposcopy the “borderline” groups like women with HPV-positive NILM could arguably be made best by reliance on shorter-term risks. Therefore, we confirmed the findings based on 3-year risk by consideration of risks of diagnosing CIN3+ within 18 months, concordant with the current U.S. recommendations to retest this group at 1 year (to which we added 6 months to accommodate delayed return). We were not able to assess directly the risk if such women were referred immediately, as this was not the practice at KPNC.

To consider potential management of these numerous HPV type group-cytology risk strata, the risk for each stratum was compared with the relevant cytology-related benchmarks used for current U.S. management guidelines.^{25,32} The benchmark risk thresholds for this population are given in the legend to Table 2 (HSIL+ = 52.0% justifying consideration of diagnostic cone if precancer is not found; LSIL = 5.8% mandating colposcopy; ASC-US = 2.3% concordant in the U.S. with 1-year return; NILM = 0.17% for 3-year return).

The 3-year risk of CIN3+ for women testing Onclarity-positive, combining all typing channels, was 7.1% (95% CI 6.8–7.4), which significantly exceeds the benchmark risk for triage to colposcopy in this KPNC population, as established by an LSIL cytologic result (5.8%). Guidelines committees in the U.S. have reasoned that it would not be practical to refer all HPV-positive women to immediate colposcopy. We examined whether cytology could be used to triage the HPV positive women and further clarify management.

When compared to benchmarks, the management of most HPV group/cytology strata was evident, with some interesting questions remaining. The risk of HPV16-positive HSIL+ was so high (60.6% risk of CIN3+ within 3 years, exceeding 80% for CIN2) that a diagnostic cone might be indicated if initial colposcopic findings failed to find precancer. HPV18/45-positive HSIL+ was lower-risk but similar. HPV18/45-positive NILM predicted a risk below the threshold for immediate colposcopy, but poorly detected glandular lesions including cancers might dictate greater concern.

The 3-year CIN3+ risks for HPV33/58/31/52 were significantly greater than those for HPV 51/35/39/68/56/59/66, for each cytology result. When benchmarked against risk-management thresholds, the common combination of HPV 51/35/39/68/56/59/66-associated ASC-US/LSIL was low (2.0, CI 1.7–2.4), in line with the risk of ASC-US cytology (without HPV triage). This risk level suggests a 1-year return in the U.S., rather than immediate colposcopy as typically

Table 3. Distribution of HPV type groups and cytologic results for HC2+ invasive cancers

HPV	Cytology	SCC (#/%)	ADC (#/%)
Total	Overall	70 (100%)	53 (100%)
	High-grade	45 (64%)	29 (55%)
	LSIL/ASC-US	15 (21%)	5 (9%)
	NILM	10 (14%)	19 (36%)
HPV16	Overall	39 (56%)	30 (57%)
	High-grade	29	10
	LSIL/ASC-US	7	3
	NILM	3	17
else HPV18/45	Overall	12 (17%)	21 (40%)
	High-grade	8	18
	LSIL/ASC-US	3	2
	NILM	1	1
else HPV31/33/58/52	Overall	12 (17%)	0 (0%)
	High-grade	5	0
	LSIL/ASC-US	3	0
	NILM	4	0
else HPV51/35/39/68/56/59/66	Overall	6 (9%)	1 (2%)
	High-grade	3	1
	LSIL/ASC-US	2	0
	NILM	1	0

Five SCC lacked Onclarity data, and one was Onclarity-negative, HC2+ NILM. Three ADC lacked typing data, one was Onclarity-negative, HC2- NILM. ADC included adenosquamous cancers.

recommended for HPV-positive ASC-US and all LSIL. Of the 1,123 women with HPV51/35/39/68/56/59/66 ASC-US/LSIL (~15% of the total of Onclarity-positive specimens), 992 attended immediate colposcopy, with the following results: 96 CIN2, 25 CIN3 and 2 squamous cell cancers. The two cancers both had LSIL cytologic results; one was HPV35/HPV51-positive and one was HPV39-positive. Important to interpretation, both were chosen for typing specifically because they had been diagnosed with cancer, not as part of the random draw of HC2-positive women.

A great majority of cancers diagnosed in the KPNC screening population were HC2-positive, thus, their corresponding specimens tended to be collected and typed during the enrollment period of the PaP cohort (*i.e.*, the tested subset was thus more likely to be HC2 positive than the full group of cancers in the cohort). Table 3 shows the enrollment HPV type/cytology test results for 123 HC2+ cancers in the Onclarity analysis. The sample included 70 SCC and

53 adenocarcinomas (including adenosquamous, while 15 others with unknown histology were excluded). Adenocarcinomas were virtually always associated with HPV16 (often NILM) or HPV18/45 (mainly high-grade cytology, about half of which was AGC rather than HSIL/cancer).

In an ancillary analysis, we divided the women tested by Onclarity into younger versus older, around the approximate median age of the cases of CIN3+ in the tested population (30–39 versus 40–65). There was a non-significant decrease in 3-year risk of CIN3 in the older group relative to the younger group,³³ possibly related to fewer HPV16 and HPV18/45 infections, but there were no striking differences in risks by age when examined within each of the HPV/cytology strata (data not shown).

Finally, we repeated the main analyses (computing 3-year risks for the hierarchical HPV type channels in Supporting Information Table 2 and the HPV/cytology risk stratification in Supporting Information Table 3) using CIN2+ as the outcome instead of CIN3+. Although the risks were higher (with HPV16+ high-grade cytology predicting >80% risk of CIN2+ diagnosis within 3 years), the risk relationships were similar. For example, the 3-year risk of CIN2+ among women with ASC-US/LSIL and HPV51/35/39/68/56/59/66 was 8.6% (95% CI 8.0–9.2) with a strong predominance of CIN2 rather than CIN3+; the risk was lower than the benchmark risk of CIN2+ that mandates immediate colposcopic referral (the risk associated with LSIL, 16.7%, 95% CI 15.7–17.7).

Discussion

The results of this analysis indicate that HPV testing with Onclarity, yielding partial typing data, can be combined with cytology triage to predict risks of cervical precancer/cancer in refined strata varying from extremely high to extremely low risk. As one possible strategy, these risk strata could be condensed into four action bands, permitting risk-based and simplified management guidelines.

About 95% of women aged 30–65 screened in the KPNC population would test HPV-negative (by HC2 or Onclarity), and the clinical value of adding cytology when HPV testing was negative was restricted to the small fraction of screened women (0.014%) with concurrent high-grade cytology results. If HPV-negative women can be considered at low risk, and rescreened in several years (the low-risk action band), the major challenge is to determine how to manage HPV-positive women. We therefore focused on the possible use of cytology to triage Onclarity-positive women following primary standalone HPV testing (not cotesting).

We opted to divide HPV types into four strata (HPV16, else HPV18/45, else 33/58/31/52, else 51/35/39/68/56/59/66. HPV18/45 are separated because of associations with difficult-to-detect glandular lesions and with cancers (particularly adenocarcinomas). On the basis of short-term prospective risk of CIN3/AIS/cancer (CIN3+), one might group them differently. The groupings analyzed here are just one possibility that proved sensible in the KPNC setting.

HPV testing requires triage to guide colposcopic referral and treatment. We used partial HPV typing and cytologic results for triage. Risk estimates from novel screening approaches like the ones generated in this analysis can be compared to “benchmark” clinical management thresholds.^{25,26} To briefly review the principles, in the U.S., the management thresholds rely on long-accepted cervical screening practice based on cytologic results without reference to HPV testing (*i.e.*, for HSIL+ including AIS and cancer, LSIL, ASC-US, NILM).³²

Each of the four risk strata defined by the thresholds corresponds to a matching clinical action.²⁵ Specifically, the very high risk following cytology of HSIL+ indicates that beyond the need for colposcopy, there is a high likelihood of precancer/cancer and probable need for a diagnostic LEEP if colposcopy does not diagnose precancer. The risk corresponding to LSIL typically lead to colposcopic biopsy to rule out precancer. ASC-US, in the absence of HPV triage, can be managed by return testing, at 1 year in the U.S. guidelines. A negative cytologic result of NILM leads to rescreening at a regular interval of 3 years in the U.S., longer in some other places.

With regard to HPV typing and its role in cervical screening, it proved possible, using benchmarking and risk-action thresholds, to combine the risk strata formed by combining HPV type-groups with cytology triage logically into these same action bands. They range from, a recommendation to consider diagnostic cone if precancer is not found, colposcopy, retest after giving HPV infections a chance to clear (Table 2), or routine screening interval (5 years is recommended in the U.S. for negative cotests, but many clinicians opt for 3 years instead).

Thus, the approach demonstrates that increased precision in management based on risk is not incompatible with simplified management guidelines. Some clinicians favoring simplicity might wish to know only the broad management recommendations for a woman, while others may want access to the details of the underlying risks and how they were generated (perhaps through comments accompanying the laboratory report). Even after a woman is referred to colposcopy, knowing risk estimates might guide the numbers of biopsies taken.³⁴

The major choice generated by partial typing with Onclarity followed by cytology triage is a difficult one, and the answer depends on societal values not scientific data. The aggregate risk faced by the sizeable group of women with HPV 51/35/39/68/56/59/66-associated ASC-US/LSIL is low, that is, about the same as that for ASC-US cytology (without HPV triage). Is the risk low enough to defer colposcopy for a year rather than refer immediately to colposcopy? If these women were retested at 1 year rather than attending immediate colposcopy as is currently recommended, we can roughly project (based on the clearance rates observed among KPNC women with the same types and NILM cytology that were retested at around 1 year) that about 1/3 to 1/2 of women

would retest as HPV-negative NILM.³⁵ Therefore, switching to 1-year follow-up for the lowest-risk women with HPV-positive ASC-US/LSIL cytology could avert colposcopy for approximately 5% of HPV-positive women, or for roughly 0.25% of all screened women. In the U.S., this could amount to tens of thousands fewer referrals annually. However, the harm would be a delay in diagnosis by 1+ years for ~3% of women with underlying CIN3+, including delay for a few cases of cancer per million women screened. Moreover, a non-negligible fraction of women with CIN2 due to these lower risk HPV types would have colposcopy deferred (96/665 or 14.4%), which might or might not be appropriate. In summary, the desirability of distinguishing HPV31/33/52/58 from HPV 51/35/39/68/56/59/66, versus combining the 11 types into a large group, hinges largely on the value ascribed to avoiding colposcopy for tens of thousands of women versus earlier possible diagnosis of a very small number of cancers, and a larger number of precancers. A cost-benefit analysis that takes into consideration the full range of societal and personal benefits and harms is necessary to guide us as we decide whether to use HPV partial typing to reduce colposcopy referrals.

An unanticipated ancillary benefit of dividing HPV-positive results into four risk groups as described might be efficient vaccine monitoring, because the seven higher-risk HPV types (HPV16/18/31/33/45/52/58) turn out to be the same ones targeted by the newly approved nonavalent HPV vaccine.³⁶ This concordance represents a kind of independent validation of both the present analysis and the estimates used to design the vaccine.

Another possible “off-label” use of more detailed HPV typing is in follow-up of individual patients. There is no full typing assay approved for clinical use by the U.S. FDA and, to our knowledge, no submission of a typing test is planned. Having even four HPV type bands can be useful when monitoring whether a given woman has a persistent or new infection, for example, post-treatment when HPV infection is found.³⁷ In passing, we note that the decision to combine HPV33 and HPV58 in the design of Onclarity might obscure the high risk of CIN3+ predicted by HPV33 in some cohort studies,^{9,38} whereas HPV58 is particularly important in some series in Asia. It also might have been useful to separate HPV35 from the lowest-risk carcinogenic types, but the effects of these design features are less important once the channels are grouped.

We noted that the 3-year risk of CIN3+ for all Onclarity-positive women exceeded the risk of LSIL cytology, for which immediate colposcopy is recommended. However, to refer all HPV-positive women on the basis of 3-year risks of CIN3+ is unlikely to prove cost-effective. Many women with early infections would not yet (or never) have detectable lesions.³⁹ This is a case where it is important to consider the time frame of risk estimates; ideally, one would use immediate and not 3-year risk to determine which HPV-positive women need immediate colposcopy. We were not able to calculate

immediate risk, because HPV-positive women with NILM, the majority of all infected women,⁴⁰ are not retested until 1 year at KPNC.

The HC2-positive invasive adenocarcinoma cases that we typed were linked almost exclusively to HPV16 and were often cytologically normal at enrollment, or were related to HPV18/45 (almost all with high-grade cytology). HPV16 and HPV18/45 adenocarcinomas are an example of convergent viral evolution; the same histopathology can be caused by two very different sets of HPV.⁴¹ The clinical significance, if any, of these viral genotype-phenotype associations is not yet clear.

The screening strategy suggested by these data from the low-risk, well-screened KPNC population should be evaluated in other settings. We urge the use of CIN3+ as the surrogate endpoint to assess cancer risk, rather than CIN2, because many cases of CIN2 are linked to HPV types not important for cancer. As we compare various screening and triage options to determine their relative effectiveness, we must

realize that due to cost, randomized clinical trials will not be possible for most comparisons. Comparative effectiveness research on the many decisions we need to make, as HPV testing and other technologies enter cervical screening, will rely mainly on large observational studies in which residual specimens are tested multiple times using different candidate tests. Screening and management guidelines currently tend to devalue non-randomized observational analyses and health decision modeling.^{3,32} Yet, it is apparent that we will need to rely increasingly on these sources of evidence in the years ahead.

In conclusion, it is possible to combine HPV partial typing in primary screening and reflex cytology to stratify risk of CIN3+ quite powerfully. The resultant detailed risk strata can be condensed into a few action bands, satisfying the criterion of “equal management of equal risk,”²⁵ and satisfying the somewhat conflicting calls for precision management and simplicity of practice guidelines.

References

- Bouvard V, Baan R, Straif K, et al. A review of human carcinogens—Part B: biological agents. *Lancet Oncol* 2009;10:321–2.
- Schiffman M, Castle PE, Jeronimo J, et al. Human papillomavirus and cervical cancer. *Lancet* 2007;370:890–907.
- 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. *CA Cancer J Clin* 2012;62:147–72.
- Gage JC, Schiffman M, Katki HA, et al. Reassurance against future risk of precancer and cancer conferred by a negative human papillomavirus test. *J Natl Cancer Inst* 2014;106:pil: dju153.
- Wentzensen N, Schiffman M, Palmer T, et al. Triage of HPV positive women in cervical cancer screening. *J Clin Virol* 2016;76(Suppl 1):S49–S55.
- Kinney W, Wright TC, Dinkelspiel HE, et al. Increased cervical cancer risk associated with screening at longer intervals. *Obstet Gynecol* 2015;125:311–5.
- Herbert A. Primary HPV testing: a proposal for co-testing in initial rounds of screening to optimise sensitivity of cervical cancer screening. *Cytology*, in press.
- 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;136:178–82.
- Schiffman M, Burk RD, Boyle S, et al. A study of genotyping for management of human papillomavirus-positive, cytology-negative cervical screening results. *J Clin Microbiol* 2015;53:52–9.
- Schiffman M, Vaughan LM, Raine-Bennett TR, et al. A study of HPV typing for the management of HPV-positive ASC-US cervical cytologic results. *Gynecol Oncol* 2015;138:573–8.
- Cuzick J, Bergeron C, von Knebel Doeberitz M, et al. New technologies and procedures for cervical cancer screening. *Vaccine* 2012;30:F107–F16.
- Castle PE, Gutierrez EC, Leitch SV, et al. Evaluation of a new DNA test for detection of carcinogenic human papillomavirus. *J Clin Microbiol* 2011;49:3029–32.
- Wright TC, Jr., Stoler MH, Agreda PM, et al. Clinical performance of the BD onclarity HPV assay using an adjudicated cohort of BD SurePath liquid-based cytology specimens. *Am J Clin Pathol* 2014;142:43–50.
- Schiffman M, Wentzensen N. A suggested approach to simplify and improve cervical screening in the United States. *J Low Genit Tract Dis* 2016;20:1–7.
- Castle PE, Shaber R, LaMere BJ, et al. Human papillomavirus (HPV) genotypes in women with cervical precancer and cancer at Kaiser Permanente Northern California. *Cancer Epidemiol Biomarkers Prev* 2011;20:946–53.
- Castle PE, Fetterman B, Poitras N, et al. Five-year experience of human papillomavirus DNA and Papanicolaou test cotesting. *Obstet Gynecol* 2009;113:595–600.
- LaMere BJ, Howell R, Fetterman B, et al. Impact of 6-month frozen storage of cervical specimens in alkaline buffer conditions on human papillomavirus genotyping. *J Virol Methods* 2008;151:298–300.
- LaMere BJ, Kornegay J, Fetterman B, et al. Human papillomavirus genotyping after denaturation of specimens for Hybrid Capture 2 testing: feasibility study for the HPV persistence and progression cohort. *J Virol Methods* 2007;146:80–5.
- Arbyn M, Snijders PJ, Meijer CJ, et al. Which high-risk HPV assays fulfil criteria for use in primary cervical cancer screening? *Clin Microbiol Infect* 2015;21:817–26.
- Rucker G, Messerer D. Remission duration: an example of interval-censored observations. *Stat Med* 1988;7:1139–45.
- Odell PM, Anderson KM, D’Agostino RB. Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics* 1992;48:951–9.
- Dorey FJ, Little RJ, Schenker N. Multiple imputation for threshold-crossing data with interval censoring. *Stat Med* 1993;12:1589–603.
- Robertson T, Wright F, Dykstra R. Order restricted inference. New York: Wiley, 1988.
- Fletcher R. Practical methods of optimization, 2nd edn. New York: Wiley, 1987.
- Katki HA, Schiffman M, Castle PE, et al. Benchmarking CIN 3+ risk as the basis for incorporating HPV and Pap cotesting into cervical screening and management guidelines. *J Low Genit Tract Dis* 2013;17:S28–S35.
- Katki HA, Kinney WK, Fetterman B, et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. *Lancet Oncol* 2011;12:663–72.
- Carreon JD, Sherman ME, Guillen D, et al. CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from. *A histological review of population-based cervical samples. Int J Gynecol Pathol* 2007;26:441–6.
- Chen Z, DeSalle R, Schiffman M, et al. Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97. *J Virol* 2009;83:1443–55.
- Guan P, Howell-Jones R, Li N, et al. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. *Int J Cancer* 2012;131:2349–59.
- Castellsague X, Diaz M, de Sanjose S, et al. Worldwide human papillomavirus etiology of cervical adenocarcinoma and its cofactors: implications for screening and prevention. *J Natl Cancer Inst* 2006;98:303–15.
- Katki HA, Schiffman M, Castle PE, et al. Five-year risks of CIN 3+ and cervical cancer among women with HPV-positive and HPV-negative high-grade Pap results. *J Low Genit Tract Dis* 2013;17:S50–S5.
- Massad LS, Einstein MH, Huh WK, et al. 2012 updated consensus guidelines for the management of abnormal cervical cancer screening tests and cancer precursors. *Obstet Gynecol* 2013;121:829–46.
- Gage JC, Katki HA, Schiffman M, et al. Age-stratified 5-year risks of cervical precancer among

- women with enrollment and newly detected HPV infection. *Int J Cancer* 2015;136:1665–71.
34. Wentzensen N, Walker JL, Gold MA, et al. Multiple biopsies and detection of cervical cancer precursors at colposcopy. *J Clin Oncol* 2015;33:83–9.
 35. Katki HA, Schiffman M, Castle PE, et al. Five-year risks of CIN 3+ and cervical cancer among women who test Pap-negative but are HPV-positive. *J Low Genit Tract Dis* 2013;17:S56–S63.
 36. Joura EA, Giuliano AR, Iversen OE, et al. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *N Engl J Med* 2015;372:711–23.
 37. Mariani L, Sandri MT, Preti M, et al. HPV-testing in follow-up of patients treated for CIN2+ lesions. *J Cancer* 2016;7:107–14.
 38. Castle PE, Stoler MH, Wright TC, Jr., et al. 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.
 39. 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.
 40. Kovacic MB, Castle PE, Herrero R, et al. Relationships of human papillomavirus type, qualitative viral load, and age with cytologic abnormality. *Cancer Res* 2006;66:10112–9.
 41. Mirabello L, Yeager M, Cullen M, et al. HPV16 sublineage associations with histology-specific cancer risk using HPV whole-genome sequences in 3200 women. *J Natl Cancer Inst*, 2016;9:108.