Human Papillomavirus Testing for Primary Screening of Cervical Cancer Precursors[1]

Samuel Ratnam, Eduardo L. Franco,[2] and Alex Ferenczy

Newfoundland Public Health Laboratory, Department of Health, Division of Community Medicine, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, A1B 3T2 Canada [S. R.]; Departments of Oncology and Epidemiology, McGill University, Montreal, Quebec H2W 1S6 Canada [E. L. F.]; Departments of Pathology and Obstetrics and Gynecology, McGill University and the Sir Mortimer B. Davis-Jewish General Hospital, Montreal, Quebec, H3T 1E2 Canada [A. F.]

Abstract

Our objective was to determine whether the addition of human papillomavirus (HPV) testing to screening cytology improves the detection of cervical cancer precursors. Women of ages 18–69 years underwent conventional Pap cytology and HPV DNA testing in a multicenter study in Newfoundland, Canada. Those with positive cytology and/or HPV and a random sample of those with dual negative results were referred for colposcopy. The study enrolled 2098 women. The relative sensitivity of HPV testing was significantly higher than cytology for all-grade squamous intraepithelial lesions [SILs; 73%; 95% confidence interval (CI), 62–82] and high grade SILs (HSILs; 90%; 95% CI, 74–97) but had lower relative specificity (62% for all-grade SILs and 51% for HSILs) than most cytological cutpoints. The rate of combined correct results for all-grade lesions was higher for HPV testing (68.8%) than for any cytological cutpoint (equivocal, 52.3%; LSILs, 51.6%; HSILs, 44.5%). The combination of HPV and an LSIL cutpoint had a negative predictive value of 68% (95% CI, 52–80) for all SILs and 100% (95% CI, 91–100) for HSILs, while referring for colposcopy only 12% of the women. We concluded that HPV testing in conjunction with cytology improved the screening efficacy of cytology alone and may allow for a more effective and safe primary screening program with increased screening intervals.

Introduction

There is compelling evidence that screening by Pap cytology has been the major contributing factor for the control of invasive cervical cancer. Descriptive epidemiological studies have clearly shown a dramatic decrease in both the incidence and mortality rates attributable to squamous cell carcinoma subsequent to the introduction of cytology screening in Scandinavian countries and in North America (1–3). Despite its successful track record, however, Pap cytology is far from perfect as a laboratory method, and in many settings, especially developing countries, cytology-based programs have failed to reduce cervical cancer rates substantially (4, 5).

There are many reasons for the pitfalls of cervical cancer screening systems, including cervical sampling errors and laboratory errors in screening and interpretation. Pap cytology is based on highly subjective interpretation of morphological alterations and is also dependent on optimally collected samples. Also, the highly repetitive nature of the work of screening many Pap smears leads to fatigue, which invariably causes errors in interpretation. In addition, false-positive cytology results lead to unnecessary and frequently invasive procedures in a fairly large number of women, which in turn result in increased patient anxiety and costs.

The high cost of screening and the resulting unnecessary follow-up procedures have led international health agencies such as the WHO and the Union Internationale Contre le Cancer to recommend increasing screening intervals from annual to every 3 years. However, extending screening intervals prompts the serious concern of interval cancers related to false-negative Pap smears that will not be revealed upon repeated screening. Such cases pose important medical, financial, and legal implications; the latter being a particularly acute problem in the United States, where false-negative Pap tests are among the most frequent reasons for medical malpractice litigation.

This state of affairs elicited interest from the medical technology industry for developing new tests with adequate sensitivity and specificity for detecting clinically significant cancer precursors. One such method is HPV[3] testing via viral DNA detection, based on the rationale that there is now consensus to regard cervical cancer and its precursors a disease caused by sexually transmitted, high-risk HPV genotypes (6). The objective of this study was thus to determine and compare the effectiveness of HPV DNA testing and cervical cytology for the detection of cervical disease in a primary screening capacity.

Materials and Methods

Study Design. Between November 1996 and August 1998 we carried out a province-wide study of women in the age group

Received 2/29/00; revised 7/7/00; accepted 7/26/00.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] This study was partially supported by grants from the Newfoundland Cancer Treatment and Research Foundation and from the Memorial University of Newfoundland. This study was approved by the Human Investigation Committee, Faculty of Medicine, Memorial University of Newfoundland, and by the institutional review boards of collaborating centers. E. L. F. is the recipient of a Distinguished Scientist Award from the Medical Research Council of Canada.
[2] To whom requests for reprints should be addressed, at Department of Oncology, McGill University, 546 Pine Avenue West, Montreal, Quebec, H2W 1S6 Canada. Phone: (514) 398-6032; Fax: (514) 398-5002; E-mail: eduardof@oncology.lan.mcgill.ca.

[3] The abbreviations used are: HPV, human papillomavirus; ASCUS, atypical squamous cells of undetermined significance; HSIL/LSIL, high/low-grade squamous intraepithelial lesion(s); CIN-I and CIN-II/III, cervical intraepithelial neoplasia grades I and II/III, respectively; HC-I/II, first/second generation Hybrid Capture; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval.

18–69 years who presented for routine Pap screening in 10 clinics representing different regions in Newfoundland, Canada. Basic demographic information was obtained at enrollment for those consenting to participate. Ecto- and endocervical specimens for cytology and HPV testing were collected from all participants using a cytobrush and an Ayre's type spatula. All women with an abnormal smear, with a positive HPV test result, or both were referred to Ob/Gyn specialists for colposcopic examination and biopsy if indicated. A separate group of women with negative cytology and negative HPV results was randomly chosen to be referred for colposcopy, corresponding to approximately 1 for every 10 cases with dual negative results. This was done to permit the calculation of sensitivity and specificity for the two screening tests (see below). Colposcopies and biopsies were carried out in accordance with standard provincial and national recommendations. Histological examination of biopsies was done at individual hospital-based laboratories.

Pap Cytology. Conventional Pap smears were processed and read in one of four regional, accredited cytology laboratories. For colposcopy referral purposes, a positive Pap smear was considered to be any abnormality ranging from ASCUS to SILs and invasive cancer. Pap smears were evaluated in accordance with the Bethesda classification system.

HPV Testing. The HPV DNA assays were carried out independently of the Pap smear screens at the Newfoundland Public Health Laboratory using the Hybrid Capture assay (Digene, Inc., Beltsville, MD). Specimens were collected and processed in accordance with the manufacturer's instructions. The first generation Hybrid Capture (HC-I) tube assay was used to test all subjects admitted until September 30, 1997. For those admitted on or after October 1, 1997, the second generation Hybrid Capture (HC-II) microtiter assay was used, also strictly according to the manufacturer's instructions. Only the high-risk HPV DNA probe mixture was used in the study. For HC-I, this probe included the high oncogenic risk HPV types 16, 18, 31, 33, 35, 45, 51, 52, and 56. The HC-II high-risk HPV mixture contained four additional HPVs, 39, 58, 59, and 68. A positive HPV test was defined as that with relative light units:positive controls ratios of 1.0 or greater, implying that the cervical specimen contained one or more of the above oncogenic HPVs.

Statistical Analysis. The study's objective was to compare the diagnostic performance of cytology and HPV testing and to assess whether the combination of the two tests provided a gain in screening efficacy with respect to that of Pap cytology alone. Sensitivity, specificity, PPVs, and NPVs and their 95% CIs were calculated using standard formulae. Calculations were performed for three different definitions (cutpoints) of positive cytology: ASCUS or worse, LSIL or worse, and HSIL. In addition, computations of indexes included two different definitions of cervical disease: CIN-I or worse and CIN-II/III or worse. The latter definitions were based on two different gold standards for ascertaining disease: a stringent one based on the histological examination alone, and a liberal one, in which the colposcopical diagnosis was considered if no histology was available.

For the combination testing, we considered a positive result when either cytology or HPV testing resulted positive. Such combination testing always yields a gain in sensitivity but a loss in specificity (relative to cytology) that is attributable to chance alone. Therefore, to measure the true improvement in screening efficacy, we compared the indexes of the combination testing against expected values obtained after correcting the sensitivity and specificity of cytology for the contribution of a hypothetical, random adjunct test that had the same positivity as HPV testing in the same population (7).

Method for Correcting for Verification Bias. Having the disease outcome determined in a random sample of women who tested negative simultaneously for Pap and HPV allowed us to derive more valid estimates of sensitivity and specificity for the two screening tests. However, to prevent the effect of verification bias, an adjustment procedure was necessary. This bias is caused by the fact that only a fraction of the women testing negative by both tests (~10%) and a much larger proportion of those with at least one test positive (100% as specified in the protocol) were referred for colposcopical verification, which decreases the relative proportions of negative subjects within the disease case and non-case series. Given the prevalence of cervical lesions and the rate of positive Pap and HPV results in our population, the bias causes a falsely elevated sensitivity and a falsely decreased specificity.

As an illustration of the impact of the verification bias, we may consider a hypothetical situation in which a single test with 80% sensitivity and 67% specificity were applied to a population of 1000 subjects with 10% disease prevalence. With complete verification of the disease status, the frequencies of true positives, true negatives, false positives, and false negatives would be: 80, 600, 300, and 20, respectively. If verification is randomly restricted to 80% of those testing positive and 10% of those testing negative, the screening table frequencies become, respectively, as above: 64, 60, 240, and 2. Using this sample to compute the above indexes yields 97% for sensitivity and 20% for specificity, values that do not represent the true performance of this hypothetical test. It is noteworthy that the PPV (21%) and NPV (97%) are not affected by the verification bias; the same estimates are obtained with either the complete or the restricted table frequencies.

We adjusted for verification bias in a dual test situation by correcting the frequencies of test results by adding to them the proportion of unverified subjects with the same test result who had the same lesion grade. This was done as follows. Denote the frequency of subjects in each combination of test results and lesion category by F_{CHL}, where the subscripts C, H, and L indicate the results for cytology and HPV, and the lesion status, respectively. Subscript values for cytology denote negative (0), ASCUS (1), LSIL (2), or HSIL+ (3) results. For HPV, there are two possible results: negative (0) and positive (1), and for lesion status: negative (0), CIN-I (1), or CIN-II/III (2). Let U_{CH} be the frequency of unverified subjects in each combination of C and H. To compensate for the verification bias the adjusted frequency for each combination was calculated as follows:

$$A_{CHL} = F_{CHL} + U_{CH} \times [F_{CHL}/(F_{CH0} + F_{CH1} + F_{CH2})]$$

To use this formula, one has to assume that for each subset of women with a given combination of cytology and HPV results, the distribution of lesion grades among those not referred for colposcopy (and thus with unverified lesion status) is the same as for those with lesion grade ascertained by colposcopy. This is a plausible assumption because the Pap and HPV test results were the only criteria determining referral for colposcopy, and thus women without colposcopic results are likely to have the same prevalence of lesion categories as those in the same stratum represented by the combination of Pap and HPV result. The same assumption is not tenable for histological verification because the decision to biopsy is conditional also on the clinician's impression during the colposcopic examination.

Any estimates of sensitivity and specificity that were

Table 1 Distribution of screening cytology and HPV testing results by selected characteristics

Category (no. of subjects)	Positivity by Pap cytology		HPV Positivity (%)
	ASCUS (%)	Any grade SIL (%)	
Age group			
<25 (401)	17 (4.1)	31 (7.5)	69 (16.7)
25–34 (1098)	62 (5.6)	45 (4.0)	130 (11.7)
35–44 (536)	21 (3.9)	15 (2.8)	27 (5.0)
45+ (59)	1 (3.6)	1 (3.6)	1 (3.6)
Clinic			
Churchill Square Clinic (503)	16 (3.2)	9 (1.8)	41 (8.2)
Gander Medical Clinic (190)	3 (1.6)	9 (4.7)	15 (7.9)
Family Medicine Clinic (441)	10 (2.3)	17 (3.9)	29 (6.6)
Murray Clinic (174)	9 (5.2)	0 (0)	14 (8.0)
Grace Ambulatory Clinic (173)	26 (15.0)	33 (19.1)	48 (27.7)
Northeast Medical Clinic (199)	11 (5.5)	2 (1.0)	20 (10.1)
Ivor Smrz Clinic (225)	14 (6.2)	9 (4.0)	27 (12.0)
Other (193)	12 (6.2)	13 (6.7)	33 (17.1)
Laboratory			
St. John's (783)	50 (6.4)	55 (7.0)	118 (15.1)
Gander (387)	14 (3.6)	11 (2.8)	35 (9.0)
Grand Falls-Windsor (443)	10 (2.3)	17 (3.8)	29 (6.5)
Corner Brook (485)	27 (5.6)	9 (1.9)	45 (9.3)

Table 2 Distribution of Pap cytology and HPV test screening results by cervical lesion as ascertained by histology alone or by a combination of colposcopy and histology

Screening test	Screening test result	Lesion grade (%)[a]		
		Negative	CIN-I	CIN-II/III+
A. Lesions ascertained by histology alone				
Pap cytology	Negative	28 (43.8)	22 (34.4)	14 (21.9)
	ASCUS	14 (48.3)	11 (37.9)	4 (13.8)
	LSIL	7 (30.4)	11 (47.8)	5 (21.7)
	HSIL	4 (33.3)	1 (8.3)	7 (58.3)
HPV test	Negative	33 (62.3)	17 (32.1)	3 (5.7)
	Positive	20 (26.7)	28 (37.3)	27 (36.0)
B. Lesions ascertained by histology and colposcopy				
Pap cytology	Negative	153 (62.2)	78 (31.7)	15 (6.1)
	ASCUS	48 (63.2)	22 (28.9)	6 (7.9)
	LSIL	25 (41.0)	30 (49.2)	6 (9.8)
	HSIL	13 (52.0)	5 (20.0)	7 (28.0)
HPV test	Negative	158 (71.2)	59 (26.6)	5 (2.3)
	Positive	81 (43.6)	76 (40.9)	28 (15.1)

[a] In A, only histologically confirmed biopsies were considered ($n = 128$); in B, the colposcopical diagnosis was used if no histological information was available ($n = 408$).

based on the actual frequency tables cross-classifying test result and presence of disease without the above correction were considered to be relative estimates for the purposes of test comparisons.

Results

A total of 2098 women participated in this study, corresponding to a response rate of 65%. The subjects' mean age was 30 years (median, 30; 1st quartile, 26; 3rd quartile, 35 years). A total of 193 (9.2%) women had ASCUS or worse smears by cytology, and 227 (10.8%) had positive HPV tests. Table 1 shows the frequency distribution of screening cytology results and HPV positivity by selected characteristics. In general, HPV and lesion rates decreased with age but there was considerable heterogeneity among the 10 clinics concerning the prevalence of HPV and lesions. Likewise, laboratories varied somewhat in terms of lesion and HPV prevalence, with the St. John's laboratory having the highest HPV and lesion prevalence of all four testing sites. Just over 69% of the population was tested by HC-I and 31% with HC-II.

Colposcopical examinations were performed in 263 (80.4%) of the 327 women who were positive either by cytology, HPV, or both tests. Physicians' compliance with the protocol requirement for colposcopy referrals was comparable between women with only a positive cytology (77%) and those with only a positive HPV test (75.4%) but having both tests positive was associated with a higher rate of colposcopical examinations (91.4%). An additional 145 colposcopical examinations were completed for women with no reasons for referral, corresponding to a random sample of 8.2% of those whose cytology and HPV test had been both negative. A histological diagnosis was available for 128 of the 408 colposcopies performed (31.4%). Not surprisingly, a decision to biopsy was clearly associated with the reason for referral. Among those undergoing colposcopy, 40.3% of the women with either test being positive had a histological diagnosis, whereas only 15.2% of those with negative results for both Pap and HPV were biopsied. However, women with a positive Pap were just as likely to be biopsied (39.5%) as those with a positive HPV result (40.3%).

Table 2 shows the distribution of cytology and HPV results according to the final lesion diagnosis either by histology alone or histology complemented with the colposcopical result. Both subsets displayed the same trends of increasing likelihood of a higher lesion grade given a worse cytological diagnosis or HPV positivity. There was one woman found to have an invasive squamous cell carcinoma, and she was positive both by cytology (scored as HSIL) and by HPV.

Table 3 shows the relative diagnostic indexes based on the histological diagnoses alone. HPV testing had a significantly greater relative sensitivity than cytology in any of the three severity cutpoints (ASCUS, LSIL, or HSIL) for detecting CIN of any grade as well as CIN-II/III or worse lesions. Expectedly, the relatively greater sensitivity of HPV testing was accompanied by appreciably lower relative specificity than most severity cutpoints for cytology. Overall, the rate of combined correct diagnoses (true-positive plus true-negative frequencies) for any grade lesions was higher for HPV testing (68.8%; 95% CI, 60.3–76.1) than for any cytological cutpoint (ASCUS, 52.3%; LSIL, 51.6%; HSIL, 44.5%, with non-overlapping 95% CIs with respect to the HPV test). However, the accuracy rate for HPV testing was relatively lower for CIN-II/III or worse lesions (60.2%; 95% CI, 51.5–68.2) than for most cytological cutpoints: ASCUS, 51.6%; LSIL, 68.0%; and HSIL, 78.1%. The PPV and NPV for HPV testing were greater than those for all cutpoints of cytology for detecting any grade of cervical disease. For CIN-II/III or worse lesions, however, the PPV of HPV testing was comparable with that of cytology with an LSIL cutpoint.

Table 4 shows the effect of combining HPV testing to cytology as a screening tool using histology as the only gold standard. In addition to the relative index estimates and associated 95% CIs, the expected value for each parameter is also shown, assuming the addition to the respective cytological cutpoint of a random adjunct test that mimicked the contribution by HPV testing by producing the same rate of positivity in

Table 3 Relative diagnostic indexes for Pap cytology and HPV testing in primary screening of cervical lesions ascertained by histology alone[a]

Screening test	Definition of positive result	Diagnostic index and 95% confidence interval[b]			
		Relative sensitivity	Relative specificity	PPV	NPV
A. Cervical lesion diagnosed: CIN-I or worse					
Pap cytology	ASCUS or worse	52.0 (40.9–62.9)	52.8 (39.7–65.6)	60.9 (48.7–71.9)	43.8 (32.3–55.9)
	LSIL or worse	32.0 (22.5–43.2)	79.2 (66.5–88.0)	68.6 (52.0–81.4)	45.2 (35.4–55.3)
	HSIL or worse	10.7 (5.5–19.7)	92.5 (82.1–97.0)	66.7 (39.1–86.2)	42.2 (33.6–51.3)
HPV test	Positive	73.3 (62.4–82.0)	62.3 (48.8–74.1)	73.3 (62.4–82.0)	62.3 (48.8–74.1)
B. Cervical lesion diagnosed: CIN-II/III or worse					
Pap cytology	ASCUS or worse	53.3 (36.1–69.8)	51.0 (41.3–60.7)	25.0 (16.0–36.8)	78.1 (66.6–86.5)
	LSIL or worse	40.0 (24.6–57.7)	76.5 (67.2–83.8)	34.3 (20.8–50.8)	80.6 (71.5–87.4)
	HSIL or worse	23.3 (11.8–40.9)	94.9 (88.6–97.8)	58.3 (32.0–80.7)	80.2 (72.0–86.4)
HPV test	Positive	90.0 (74.4–96.5)	51.0 (41.3–60.7)	36.0 (26.1–47.3)	94.3 (84.6–98.1)

[a] Exclusively histology used to establish cervical pathology diagnosis.
[b] Relative to lesion grade diagnosed: any CIN or worse in A and only CIN-II/III or worse in B.

Table 4 Relative diagnostic indexes for the combination of Pap cytology and HPV testing in primary screening of cervical lesions ascertained by histology alone[a]

Definition of positive result:	Diagnostic index and 95% confidence interval[b]			
	Relative Sensitivity	Relative Specificity	PPV	NPV
A. CIN-I or worse lesions				
HPV+ or Pap ≥ ASCUS	93.3 (85.3–97.1) [80.1]	32.1 (21.1–45.5) [21.9]	66.0 (56.6–74.4)	77.3 (56.6–89.9)
HPV+ or Pap ≥ LSIL	82.7 (72.6–89.6) [71.8]	50.9 (37.9–63.9) [32.8]	70.5 (60.2–79.0)	67.5 (52.0–79.9)
HPV+ or Pap ≥ HSIL	76.0 (65.2–84.2) [63.0]	58.5 (45.1–70.7) [38.3]	72.2 (61.4–80.8)	63.3 (49.3–75.3)
B. CIN-II/III or worse lesions				
HPV+ or Pap ≥ ASCUS	100 (88.6–100) [80.7]	22.4 (15.3–31.7) [21.1]	28.3 (20.6–37.5)	100 (85.1–100)
HPV+ or Pap ≥ LSIL	100 (88.6–100) [75.2]	40.8 (31.6–50.7) [31.7]	34.1 (25.0–44.5)	100 (91.2–100)
HPV+ or Pap ≥ HSIL	96.7 (83.3–99.4) [68.2]	49.0 (39.3–58.7) [39.3]	36.7 (26.9–47.7)	98.0 (89.3–99.6)

[a] Only histologically confirmed biopsies were used to diagnose lesion grade.
[b] Relative to lesion grade diagnosed: any CIN or worse in A and only CIN-II/III or worse in B. 95% confidence interval in parentheses and expected value in brackets.

this sample as HPV testing did, *i.e.*, 58.6%. Adding HPV testing to cytology increased the relative sensitivity of all its severity cutpoints significantly for any grade lesions or CIN-II/III because all of the expected values were lower than the 95% CIs for the estimates. The test combination of cytology at an LSIL cutpoint and HPV positivity detected most cases (83%) of CIN and all CIN-II/III cases. In general, addition of HPV testing to cytology also avoided the loss in relative specificity that was expected when adding a second (random) adjunct test, although the differences were significant only for LSIL and HSIL as cutpoints for diagnosing any grade lesions. NPVs were >95% for HSIL in all combinations at the expense of a <50% PPV.

Table 5 shows the diagnostic indexes for cytology, HPV, and their combination using the diagnostic definition based on histology supplemented by the colposcopical result. Two sets of results are shown: with and without correction for colposcopy verification bias. The uncorrected (relative) estimates for individual tests did not differ substantially from those calculated using histology alone (Table 3), the relative sensitivity of HPV testing for detecting both any-grade lesions or CIN-II/III was greater than that of any cytology cutpoint, and its relative specificity was comparable with the ASCUS criterion of positivity.

The comparison between uncorrected and corrected estimates indicates the extent of the verification bias attributable to the different proportions of women examined by colposcopy, depending on the results of their screening test results. For all combinations, uncorrected sensitivity estimates were on average three times greater than corrected ones for detecting any grade CIN and 25–45% greater than corrected values when CIN-II/III was used as end point. On the other hand, uncorrected specificity estimates were 5–55% lower than the equivalent ones after correction for the differential rates of colposcopic verification, with the extent of the bias being similar for both any grade CIN and CIN-II/III as end points. The magnitude of the difference between corrected and uncorrected PPV and NPV estimates (range, 5–17%, depending on the test and lesion combination) was much less pronounced because these indexes are not directly affected by the verification bias.

Discussion

Before we address the implications of this study, it is important that we consider its limitations. As is the situation with most epidemiological studies attempting to reach a population spread over a large area with a high proportion of rural residents, our response rate was not very high at 65%, which could prompt a concern about the generalizability of our findings. We have no reason to believe, however, that a selection bias would have

Table 5 Crude and corrected diagnostic indexes for cytology alone, HPV testing, and their combination in primary screening of cervical lesions ascertained by colposcopy and histology[a]

Screening test	Definition of positive result	Uncorrected (relative) estimates[b]				Corrected estimates[b]			
		Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV
A. CIN-I or worse lesions									
Pap cytology	ASCUS or worse	45.0	64.0	46.9	62.2	14.2	92.9	45.2	72.4
	LSIL or worse	28.4	84.1	55.8	62.4	8.4	97.3	56.0	72.0
	HSIL or worse	7.1	94.6	48.0	59.0	2.0	99.1	48.3	71.0
HPV test	Positive	62.1	66.1	56.5	71.2	20.8	93.3	56.3	74.0
Combination	HPV+ or ≥ ASCUS	78.1	45.2	50.2	74.5	26.3	88.9	49.4	74.5
	HPV+ or ≥ LSIL	70.4	59.8	55.3	74.1	23.3	92.2	55.4	74.4
	HPV+ or ≥ HSIL	63.9	63.6	55.4	71.4	21.3	92.9	55.4	74.1
B. CIN-II/III or worse lesions									
Pap cytology	ASCUS or worse	55.9	61.8	11.7	93.9	40.2	91.6	10.7	98.4
	LSIL or worse	38.2	80.5	15.1	93.5	26.8	96.2	15.0	98.1
	HSIL or worse	20.6	95.2	28.0	93.0	14.2	99.1	28.1	97.9
HPV test	Positive	85.3	58.0	15.6	97.7	68.1	90.6	15.4	99.1
Combination	HPV+ or ≥ ASCUS	97.1	38.5	12.5	99.3	76.3	85.9	12.0	99.3
	HPV+ or ≥ LSIL	97.1	51.3	15.3	99.5	76.3	89.3	15.2	99.3
	HPV+ or ≥ HSIL	91.2	56.1	15.9	98.6	72.0	90.3	15.7	99.2

[a] The colposcopical diagnosis was used if no histological information was available.
[b] Relative to lesion grade diagnosed: any CIN or worse in A and only CIN-II/III or worse in B. Corrected estimates control for the verification bias caused by preferential colposcopy referral of positive cases.

influenced our results because the screening performance of the two tests are independent of the women's reasons for refusing to participate. This decision was taken before screening test results were made available; therefore, it is highly unlikely that lesion rates would have varied materially between participants and nonparticipants. Even if this had occurred, it is likely that the PPV and NPV would be the affected parameters because of their dependence on prevalence, not sensitivity and specificity.

In addition, as is the case with most investigations comparing screening modalities, the gold standard for disease could not be obtained for all subjects under investigation because of the impossibility of referring for colposcopy and eventually biopsy all cases presumed free of disease by the screening tests. Although our protocol established that participating physicians should have favored colposcopic-guided histological confirmation whenever either of the screening tests was positive, colposcopies were performed in only 80% of the referred cases and, among these, biopsies were taken in 40%. Although colposcopy is not considered to be a diagnostic gold standard, it performs well in predicting invasive cancer and its high-grade precursors (8). Our colposcopists were clearly influenced by the screening results in deciding whether to biopsy, with almost three times as many women being biopsied when either screening result was positive than when none was. However, the particular combination of cytology and HPV test results did not seem to influence the biopsy rate: 40% for Pap+/HPV−, 42% for Pap−/HPV+, and 39% for Pap+/HPV+. Our estimates of screening performance would have been severely biased if colposcopy and biopsy rates had been differential with respect to the screening test with a positive result. This did not happen because physicians were equally likely to refer for colposcopy (and colposcopists were equally likely to perform a biopsy) Pap+/HPV− subjects as they were Pap−/HPV+ ones.

Our choice of a random, "control" group of women referred for colposcopy with both tests being negative enabled us to compute absolute estimates of sensitivity and specificity for Pap and HPV. However, because of the different rates of colposcopy referral given the joint test results, our estimates of screening performance had to be corrected for verification bias. Referral rates were 91% if both tests were positive, 75–77% if either the HPV or the Pap test was positive, and 8% if both tests resulted negative. Such verification bias could be eliminated by adjusting the diagnostic indexes on the basis of the expected lesion distribution among those with missing colposcopical diagnoses. Because physicians were also influenced by the knowledge of the cytological lesion grade, we adjusted the frequencies for each combination of cytological cutpoint and HPV test status to ensure improved control of the verification bias.

As expected, the corrected sensitivity estimates were invariably lower than the uncorrected ones for all combinations, whereas the corrected specificity estimates were higher than the respective unadjusted values. Although inferences about differences in test performance could be made with the relative, uncorrected values, the appropriate assessment of the anticipated screening efficacy of the two tests and their combination can only be made using the bias-free estimates. An important caveat is the fact that the correction procedure is not valid for the histological diagnoses because the colposcopist's decision to biopsy is based not only on the screening test results but also on the visual impression of the cervix during the colposcopy. However, given that the results with histology-only and histology-colposcopy diagnoses were comparable in most test and lesion combinations, it is plausible to assume that our corrected estimates using the augmented (by colposcopy) diagnoses may have reflected the true screening performance of Pap and HPV tests had histological diagnoses been available for our entire sample. The fact that our study sample was approximately representative of all cervical cancer screening activity in the province supports the generalizability of our findings to the female population of Newfoundland, which experiences higher rates of invasive cervical cancer than the average for Canada (9).

As an individual screening test to detect cervical cancer precursors, HPV testing performed relatively well compared with cytology. The sensitivity of HPV testing was greater than that of cytology in any of the severity cutpoints, regardless of the lesion grade being detected and of the disease ascertainment approach. The higher sensitivity of HPV testing was penalized by a higher false-positive rate than cytology, particularly when the latter had positivity defined at the LSIL cutpoint or higher. This was especially true for the detection of CIN-II/III or worse lesions. Overall, HPV testing was more accurate than any cytology cutpoint at detecting or ruling out any grade lesions but it performed worse than cytology for diagnosing CIN-II/III, regardless of the gold standard for lesions.

The addition of HPV testing to cytology produced substantial gains in sensitivity without imposing a large penalty on specificity in all combinations of lesion grade and diagnostic approach. Using histologically confirmed lesions, the LSIL/HPV+ combination detected all cases of CIN-II/III while missing only 17% of all-grade CIN. Although this was obtained at the expense of false-positive rates of 49% for all-grade CIN and 59% for CIN-II/III, the loss in specificity was significantly lower than that expected by chance if cytology had been augmented by a non-informative test. In population-based terms (bias-corrected estimates), a dual negative result using the combination of LSIL/HPV+ would imply virtual certainty in correctly ruling out an CIN-II/III (99%) and 74% for ruling out all-grade lesions.

Another approach for assessing the potential impact of combination testing in primary screening is by gauging sensitivity against the proportion of women who would have been referred for colposcopy because of a positive result. The latter is the best indicator of the overall burden that the screening program would place in the public health system. An economically liberal practice that has gained acceptance in North America is to refer for colposcopy all ASCUS or worse abnormalities, which in our population accounted for 9.2% of all women. This management approach would have resulted in an appallingly low sensitivity of 40% (56% before correction) for CIN-II/III. A combination testing policy of referring any LSIL or worse abnormalities or HPV-positive cases would have required 12.3% of colposcopical examinations (an extra 3% of referrals), but the sensitivity for detecting CIN-II/III would have increased to 76% (97% if uncorrected).

The above findings suggest thus that one is less likely to miss clinically significant lesions when both cytology and HPV results are simultaneously negative in women attending a screening program. If our findings are corroborated by additional studies in other populations, double testing may indeed allow screening intervals to be substantially longer than currently recommended, while keeping an acceptable margin of safety against the development of cervical cancer. It has been speculated that patients who have been negative for both high-risk HPV types and cytology have a greatly decreased risk for developing HSIL as compared with those who had negative smears alone (10–12). The expense of a double screening approach can be absorbed by the reduction in cost associated with increased screening intervals (11, 13, 14). Such a reduction in cost would also offset the extra expense of referring for colposcopy women with latent HPV infection (HPV-positive test in cytologically negative women). Indeed, Cuzick and Sasieni (11) have calculated that if the screening interval for women over the age of 30 years could be extended from every 3 years to 5 years, ~30 million pounds/year could be saved in the United Kingdom screening program, which is currently estimated to cost ~130 million pounds/year.

HPV testing in a screening capacity should only be used in women after the age of 30 years (11, 15). Indeed, at this age, the prevalence of latent HPV infections including those with high-risk oncogenic types is low (15, 16). Primary screening via HPV testing in young women would lead to detection of a high number of cases either without lesions or with low-grade lesions with high spontaneous regression rates (15). On the other hand, those of ages between 30 and 50 years tend to have persistent, high-risk HPV type infections with progression potential to high-grade lesions (11, 15, 17–19) whose incidence peaks between ages 30 and 35 years. As an illustration, in the present study the bias-corrected sensitivity of HPV testing to detect CIN-II/III was 62% among women <30 years and 82% among women ≥30 years. Similarly, the bias-corrected specificity was 87 and 94% for the same age strata, respectively.

In conclusion, the results of the present study indicate that HPV testing can have an important role in primary screening for cervical cancer either as an adjunct tool to augment existing Pap cytology programs or as a standalone test. The hybrid capture technology for detecting HPV DNA has been proven to be reproducible, is simple to perform at relatively low cost, and does not require the extensive training that conventional cytological screeners must be subjected to. In addition, it is expected that HPV testing will eventually be automated and performed by relatively unskilled personnel. Data on the role of HPV testing using self-obtained cervicovaginal samples in screening for cervical carcinoma and its precursors have recently become available, a feature which may improve compliance over current screening program (20, 21). These conditions make it a serious contender for use in a screening program in the third world (21, 22), wherein lies the heaviest burden of cervical cancer incidence and mortality.

Acknowledgments

We are indebted to Debbie Mason for assistance in data management, to Elisabeth Oates for technical assistance, and to the other members of the Newfoundland HPV Study Team. We are also indebted to Professor Stephen Walter from McMaster University, Hamilton, Ontario, Canada, and to Dr. Petra Macaskill, University of Sydney, New South Wales, Sydney, Australia, for important suggestions concerning the statistical analysis.

References

1. Miller, A. B., Anderson, G., Brisson, J., Laidlaw, J., Lepitre, N., Malcolmson, P., Mirwaldt, P., Stuart, G., and Sullivan, W. Report of a National Workshop on screening for cancer of the cervix. Can. Med. Assoc. J., *145:* 1301–1327, 1991.

2. Parkin, D. M., Nguyen-Ding, X., and Day, N. E. The impact of screening on the incidence of cervical cancer in England and Wales. Br. J. Obstet. Gynaecol., *92:* 150–157, 1985.

3. Hakama, M., Chamberlain, J., Day, N. E., Miller, A. B., and Prorok, P. C. Evaluation of screening programs for gynaecological cancer. Br. J. Cancer, *52:* 669–673, 1985.

4. Koss, L. G. The Papanicolaou test for cervical cancer detection: a triumph and a tragedy. J. Am. Med. Assoc., *261:* 737–743, 1989.

5. McGoogan, E. Advantages and limitations of automated screening systems in developing and developed countries. *In:* E. Franco and J. Monsonego (eds.), New Developments in Cervical Cancer Screening and Prevention, pp. 317–322. Oxford, United Kingdom: Blackwell Scientific Publications, Ltd., 1997.

6. IARC Working Group. IARC Monographs on the Evaluation of Carcinogenic Risks. Human Papillomavirus. IARC Scientific Publications, Vol. 64, pp. 277–282. Lyon, France: IARC, 1995.

7. Franco, E. L., and Ferenczy, A. Assessing gains in diagnostic utility when human papillomavirus testing is used as an adjunct to Papanicolaou smear in the triage of women with cervical cytologic abnormalities. Am. J. Obstet. Gynecol., *181:* 382–386, 1999.

8. Mitchell, M. F., Tortolero-Luna, G., Cook, E., Whittaker, L., Rhodes-Morris, H., and Silva, E. A randomized clinical trial of cryotherapy, laser vaporization, and loop electrosurgical excision for treatment of squamous intraepithelial lesions of the cervix. Obstet. Gynecol., *92:* 737–744, 1998.

9. National Cancer Institute of Canada. Canadian Cancer Statistics 1998. Toronto, Canada, 1998.

10. van Oortmarsen, G. J., Habema, J. D., and van Ballegooijen, M. Predicting mortality from cervical cancer after negative smear test results. Br. Med. J., *305:* 449–451, 1992.

11. Cuzick, J., and Sasieni, P. Estimates of the cost impact of introducing human papillomavirus testing into a cervical screening programme. *In:* E. Franco and J. Monsonego (eds.), New Developments in Cervical Cancer Screening and Prevention, pp. 364–372. Oxford, United Kingdom: Blackwell Scientific Publications, Ltd., 1997.

12. Bollen, L. J. M., Tjong-A-Hung, S. P., van der Velden, J., Mol, B. W., ten Kate, F. W. J., ter Schegget, J., and Bleker, O. P. Prediction of recurrent and residual cervical dysplasia by human papillomavirus detection among patients with abnormal cytology. Gynecol. Oncol., *72:* 199–201, 1999.

13. van Ballegooijen, M., Koopmanschap, M. A., and Habbema, J. D. F. The management of cervical intra-epithelial neoplasia (CIN): extensiveness and costs in the Netherlands. Eur. J. Cancer, *31A:* 1672–1676, 1995.

14. Cuzick, J., Szarewski, A., Terry, G., Ho, L., Hanby, A., Maddox, P., Anderson, M., Kocjan, G., Steele, S. T., and Guillebaud, J. Human papillomavirus testing in primary cervical screening. Lancet, *345:* 1533–1536, 1995.

15. Meijer, C. J. L., Rozendaal, L., van der Linden, J. C., Helmerhorst, T. J. M., Voorhorst, F. J., and Walboomers, J. M. M. Human papillomavirus testing for primary cervical cancer screening. *In:* E. Franco and J. Monsonego (eds.), New Developments in Cervical Cancer Screening and Prevention, pp. 338–347. Oxford, United Kingdom: Blackwell Scientific Publications, Ltd., 1997.

16. Franco, E. L., Villa, L. L., Ruiz, A., and Costa, M. C. Transmission of cervical human papillomavirus infection by sexual activity: differences between low and high oncogenic risk types. J. Infect. Dis., *172:* 756–763, 1995.

17. Ho, G. Y., Buck, R. D., Klein, S., Kadish, A. S., Chang, C. J., Palan, P., Basu, J., Tachezy, R., Lewis, R., and Romney, S. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. J. Natl. Cancer Inst., *87:* 1365–1371, 1995.

18. Gaarenstroom, K. N., Melkert, P., Walboomers, J. M., van den Brule, A. J., Vanbommel, P. F., Meyer, C. J., Voorhorst, F. J., Kenemans, P., and Helmerhorst, T. J. Human papillomavirus DNA and genotypes: prognostic factors for progression of cervical intraepithelial neoplasia. Int. J. Gynecol. Cancer, *4:* 73–78, 1994.

19. Remmink, A., Walboomers, J. M. M., Helmerhorst, T. H. J. M., Voorhorst, F. J., Rozendaal, L., Risse, E. K. J., Meijer, C. J. L. M., and Kenemans, P. The presence of persistent high risk HPV genotypes in dysplastic cervical lesions is associated with progressive disease: natural history up to 36 months. Int. J. Cancer, *61:* 306–311, 1995.

20. Sellors, J. W., Lorincz, A. T., Mahony, J. B., Mielzynska, I., Lytwyn, A., and Chernesky, M. A. Self-sampling for sexually transmitted human papillomavirus (HPV). Third EUROGIN International Congress, Paris, France, March 24–27, 1997. Abstract SS9-5, p. 126.

21. Wright, T. C., Jr., Denny, L., Kuhn, L., Pollack, A., and Lorincz, A. HPV DNA testing of self-collected vaginal samples compared with cytologic screening to detect cervical cancer. J. Am. Med. Assoc., *283:* 81–86, 2000.

22. Schiffman, M., Herrero, R., Hildesheim, A., Sherman, M. E., Bratti, M., Wacholder, S., Alfaro, M., Hutchinson, M., Morales, J., Greenberg, M. D., and Lorincz, A. T. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. J. Am. Med. Assoc., *283:* 87–93, 2000.

Human Papillomavirus Testing for Primary Screening of Cervical Cancer Precursors

Samuel Ratnam, Eduardo L. Franco and Alex Ferenczy

Cancer Epidemiol Biomarkers Prev 2000;9:945-951.