

## Interlaboratory Reliability of Hybrid Capture 2

Philip E. Castle, PhD,<sup>1</sup> Cosette M. Wheeler, PhD,<sup>2</sup> Diane Solomon, MD,<sup>3</sup> Mark Schiffman, MD,<sup>1</sup> and Cheri L. Peyton,<sup>2</sup> for the ALTS Group\*

**Key Words:** Human papillomavirus; HPV; Cervical cancer; Hybrid Capture 2; Polymerase chain reaction; PCR; Reproducibility

DOI: 10.1309/BA43HMCAJ26VWQH3

### Abstract

We evaluated the interlaboratory reproducibility of the Hybrid Capture 2 (HC2; Digene, Gaithersburg, MD), a test for oncogenic human papillomavirus (HPV) DNA, using data from 4 clinical center (CC) laboratories and the quality control (QC) laboratory participating in the ASCUS (atypical squamous cells of undetermined significance) and LSIL (low-grade squamous intraepithelial lesion) Triage Study (ALTS). Residual liquid cytology specimens were tested routinely throughout the duration of ALTS at CC laboratories, and a stratified (by time in the study) random sample of specimens was retested by the HPV QC laboratory using equivalent protocols. Of the specimens selected ( $N = 1,175$ , 5.50% of all specimens obtained), 1,072 (91.23%) had sufficient specimen volume for retesting. The  $\kappa$  value between all CC laboratories and the HPV QC laboratory was 0.84 (95% confidence interval, 0.78-0.89), with  $\kappa$  values for individual CCs and the HPV QC laboratory ranging from 0.79 to 0.89. Agreement between test results was lowest among results for women with negative cytologic findings (0.73); among those with equivocal or abnormal cytologic findings,  $\kappa$  values were 0.80 or more. These data show that HC2 is a reliable test for detecting clinically relevant oncogenic HPV DNA.

Hybrid Capture 2 (Digene, Gaithersburg, MD) using probe set B (henceforth referred to as HC2) is a DNA test for detection of cancer-associated (oncogenic) human papillomavirus (HPV) types in cervical Papanicolaou (Pap) specimens. HC2 is a sensitive test for identifying women with high-grade cervical neoplasia in screening populations<sup>1,2</sup> and for triaging women with equivocal cytologic abnormalities (atypical squamous cells [ASC]).<sup>3,4</sup> Consequently, HPV testing is being introduced into US cervical cancer prevention programs.<sup>5</sup>

In addition to good clinical performance, 2 hallmarks of clinical usefulness are test reproducibility and its adaptability to widespread use in clinical laboratories. However, systematic evaluations of HC2 reliability are lacking. As part of the ALTS<sup>6</sup> (ASCUS [atypical squamous cells of undetermined significance] and LSIL [low-grade squamous intraepithelial lesion] Triage Study), a randomized clinical trial to evaluate 3 management strategies for women with equivocal or mildly abnormal cytologic findings, HC2 HPV DNA testing was performed on residual cervical specimens at each of the 4 clinical centers. The HPV quality control (QC) laboratory retested 5% of the available specimens randomly sampled on a quarterly basis throughout the trial. We, therefore, took this opportunity to study the reliability of HC2 by evaluating the interlaboratory reproducibility between the 4 clinical laboratories and the HPV QC laboratory participating in ALTS.

### Materials and Methods

#### Study Design and Population

ALTS was a randomized clinical trial comparing 3 management strategies for women with ASCUS or LSIL: immediate

colposcopy, HPV triage, and conservative management, the latter based on a program of repeated cytologic examination. Details of this study have been published.<sup>4,6,7</sup> Briefly, women with ASCUS or LSIL cytologic findings were recruited to participate in the study at 4 clinical centers: University of Alabama at Birmingham; Magee-Women's Hospital of the University of Pittsburgh Medical Center Health System, Pittsburgh, PA; the University of Oklahoma, Oklahoma City; and the University of Washington, Seattle. National Cancer Institute and local institutional review boards approved the study.

A total of 5,060 women enrolled in the study from January 1997 to December 1998: 3,488 women with ASCUS and 1,572 with LSIL cytologic findings. Routine follow-up and exit visits concluded in January 2001. After determining eligibility and obtaining written informed consent, participants were assigned randomly by referral stratum (ASCUS or LSIL) to 1 of the 3 management arms: immediate colposcopy (referral to colposcopy regardless of enrollment test results), HPV triage (referral to colposcopy if the enrollment HPV result was positive or missing or if the enrollment cytologic result was high-grade squamous intraepithelial lesion [HSIL]), and conservative management (referral to colposcopy if the cytologic result at enrollment or follow-up was HSIL).

At enrollment, women in each arm received the same enrollment pelvic examination with collection of 2 cervical specimens, the first in PreservCyt for ThinPrep cytologic examination (Cytoc, Boxborough, MA) and HC2 testing and the second in specimen transport medium (STM; Digene) for HPV DNA typing by polymerase chain reaction (PCR). Patient referral to colposcopy at enrollment was based on the randomization arm and enrollment test results. Women in all arms of the study were reevaluated by cytologic examination every 6 months for 2 years of follow-up. An exit examination, with colposcopy scheduled for *all* women regardless of study arm or previous procedures, was performed at 2 years. We refer readers to other references<sup>4,6,7</sup> for details on randomization, examination procedures, patient management, and laboratory and pathology methods.

### HC2 Testing

HC2 is a test for detection of 13 oncogenic HPV DNA types. HC2 relies on the formation of target HPV DNA-RNA probe heteroduplexes during the hybridization step in specimens positive for one or more oncogenic HPV types (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68) and the chemiluminescence detection of these hybrids by using an alkaline phosphatase-conjugated monoclonal antibody specific to DNA-RNA complexes with dioxetane substrate in a 96-well enzyme-linked immunosorbent assay format. Signal strengths in relative light units (RLU) were compared with 1 pg/mL of HPV type 16 DNA-positive control samples (RLU/PC). The US Food and Drug

Administration-approved 1.0 RLU/PC (~1 pg/mL) was used as the threshold for a positive result.<sup>4,6,7</sup>

After liquid-based, ThinPrep cytologic slides were prepared, a 4-mL aliquot of the residual in the PreservCyt vial was used for routine HPV DNA testing by HC2 at the clinical center (CC) laboratory for each patient visit throughout ALTS. Briefly, PreservCyt aliquots were treated per the manufacturer's instructions with one tenth the volume (0.4 mL) of sample preparation reagent, mixed by vortexing, and subjected to centrifugation at 2,900g for 15 minutes. PreservCyt medium was decanted carefully from the cell pellet, and 150  $\mu$ L of denaturation reagent (~2 volumes of STM:1 volume of denaturation reagent) was added and the mixture again was mixed by vortexing. Pellets were heated at 65°C for 15 to 45 minutes until resuspended. At this point, half (75  $\mu$ L) of the resuspended and denatured pellet, the equivalent of 2.0 mL of the original specimen, was added to the HC2 test well.

A stratified random sample by time interval in the study (every quarter of the collection year with exception of the first interval, which consisted of 9 months, from January to September 30, 1997) was selected (N = 1,175, 5.5% of specimens collected) for masked retesting by the HPV QC laboratory masked to the original CC laboratory's test result. We randomly selected 80 (first 3 intervals) or 85 (last 11 intervals) specimens per interval, with the exception that all 14 specimens available from Oklahoma during the June-September 2000 interval were selected.

From September 1997 to September 1998, a 4-mL aliquot was drawn from PreservCyt specimens by the HPV QC laboratory for processing, of which half of the processed aliquot (an equivalent of 2.0 mL) was used in the HC2 assay. PreservCyt specimens with less than 4.0 mL residual were excluded from testing. After September 1998, to decrease the numbers of specimens excluded because of insufficient volume, a 2.0-mL aliquot was drawn from PreservCyt specimens by the HPV QC laboratory for processing, of which half of the processed aliquot (an equivalent of 1.0 mL) was used in the HC2 assay. Only specimens with less than 2.0 mL residual were excluded from testing. Test results for 1 QC batch were rejected owing to internal testing problems at the HPV QC laboratory, and the specimens were retested.

A total of 1,072 specimens (91.23%) with HC2 testing by the CC laboratories were retested by the HPV QC laboratory (n = 299 before September 1998 [27.89%]; n = 773 after September 1998 [72.11%]). The results of these paired tests were used to study interlaboratory reproducibility. Specimen selection and testing are summarized in **Table 1**.

### PCR Testing

We used testing data based on L1 consensus primer PGMY09/11 PCR amplification and reverse-line blot

**Table 1**  
**Specimen Selection for Retesting by the HPV QC Laboratory\***

Batch/Quarter <sup>†</sup>	Origin of Specimens Tested: Clinical Center Laboratory				Total No. of Specimens <sup>‡</sup>	
	A	B	C	D	Tested	Selected
1/September 1997	20	14	19	25	78	80
2/December 1997	23	17	15	23	78	80
3/March 1998	16	13	10	31	70	80
4/June 1998	17	10	15	31	73	85
5/September 1998	18	0	23	39	80	85
6/December 1998	34	13	15	16	78	75
7/March 1999	18	19	17	20	74	85
8/June 1999	26	6	22	20	74	85
9/September 1999	25	14	21	23	83	85
10/December 1999	22	13	14	29	78	85
11/March 2000	24	10	20	20	74	85
12/June 2000	30	13	19	19	81	85
13/September 2000	35	13	15	9	72	85
14/December 2000	20	9	24	26	79	85
Total tested	328	164	249	331	1,072	1,175
Total selected (% tested) <sup>§</sup>	367 (89.4)	179 (91.6)	268 (92.9)	361 (91.7)	1,175 (91.23)	—
Total collected (% selected) <sup>  </sup>	6,548 (5.60)	3,647 (4.91)	4,469 (6.00)	6,689 (5.40)	21,353 (5.503)	—

\* The selection was a random sample by quarter of the year in the study from the specimens available at the biorepository at the time of selection and not based on the total collected. Some specimens were delayed in their shipment to the biorepository, leading to minor differences in the percentages of specimens from each clinical site retested by the human papillomavirus (HPV) quality control (QC) laboratory.

† Quarters of years were defined from the last day of the month of the previous quarter to the last day of the month in the current quarter. For example, batch 2 included September 30 to December 31, 1997. The first time interval was an interval from the beginning of the study (January 1997) to September 30, 1997.

‡ The differences in numbers tested and selected were the result of insufficient specimen volume.

§ Number of samples selected for Hybrid Capture 2 (see text for proprietary information) retesting.

|| Number of samples obtained during the 2-year duration of the trial.

hybridization for type-specific detection<sup>8,9</sup> on the second cervical specimen collected into STM from each patient. We classified HPV status hierarchically according to cancer risk: oncogenic<sup>10</sup> (the same 13 types targeted by HC2), nononcogenic if negative for oncogenic HPV types but positive for nononcogenic types, or negative (oncogenic > nononcogenic > negative). Thus, for this analysis, a woman who had both oncogenic and nononcogenic HPV types as defined by PCR was classified as having an oncogenic HPV infection.

**Pathology**

Clinical management was based on the CC pathologists' review of cytologic and histologic material and diagnosis. In addition, all referral slides, ThinPrep slides, and histologic slides were sent to the pathology QC group based at Johns Hopkins Hospital, Baltimore, MD, for rereview and final case definition. As the surrogate for cancer risk, we chose a priori a scientific end point of histologic cervical intraepithelial neoplasia (CIN) 3 or cervical cancer (only 7 cancers were diagnosed) as diagnosed by the pathology QC group. To define histologic outcomes of CIN 3 or more severe disease, we considered all diagnoses at enrollment, during the 2-year follow-up, and at exit, recognizing that CIN 3 or more severe disease detected within 2 years of a positive HPV test result is more likely the result of a missed prevalent case than a true incident case because a single colposcopic evaluation with biopsy and histologic evaluation is not perfectly sensitive for detection of cervical precancer and cancer.<sup>11,12</sup>

**Statistical Analyses**

Spearman rank correlations were used to compare raw RLU/PC values between the HC2 test and retest. Test results were classified as positive and negative, according to the 1.0 RLU/PC positive cut point, and κ values, percentage of agreement, and percentage of positive agreement with 95% confidence intervals were calculated. The McNemar χ<sup>2</sup> test was used to test for statistical differences in test positivity (*P* < .05) on paired tests.

Paired HPV test results were stratified by CC laboratory and by CC pathologists' evaluations of cytology (negative, ASC, LSIL, HSIL or more severe) according to the Bethesda System (n = 1,065; CC cytologic diagnoses were missing for 7 women).<sup>13</sup>

Paired HC2 test results also were categorized as CC-negative/QC-negative, CC-positive/QC-negative, CC-negative/QC-positive, and CC-positive/QC-positive based on the 1.0 RLU/PC positive threshold. Comparisons of the relevant median RLU/PC values for each test in each category and restricted to discordant HC2 test pairs (CC-positive/QC-negative vs CC-negative/QC-positive) were made using a nonparametric analysis of variance test (Kruskal-Wallis). All HC2 test pairs and discordant HC2 test pairs were compared with PCR test results (oncogenic, nononcogenic, or negative) (n = 853) and with histologic outcome (pathology QC group–diagnosed CIN 3 or cervical cancer vs disease less severe than CIN 3) by using Pearson χ<sup>2</sup> tests in the subset of specimens obtained before or concurrently with the diagnosis

of CIN 3 or cervical cancer (ie, excluding those obtained after treatment) ( $n = 967$ ).

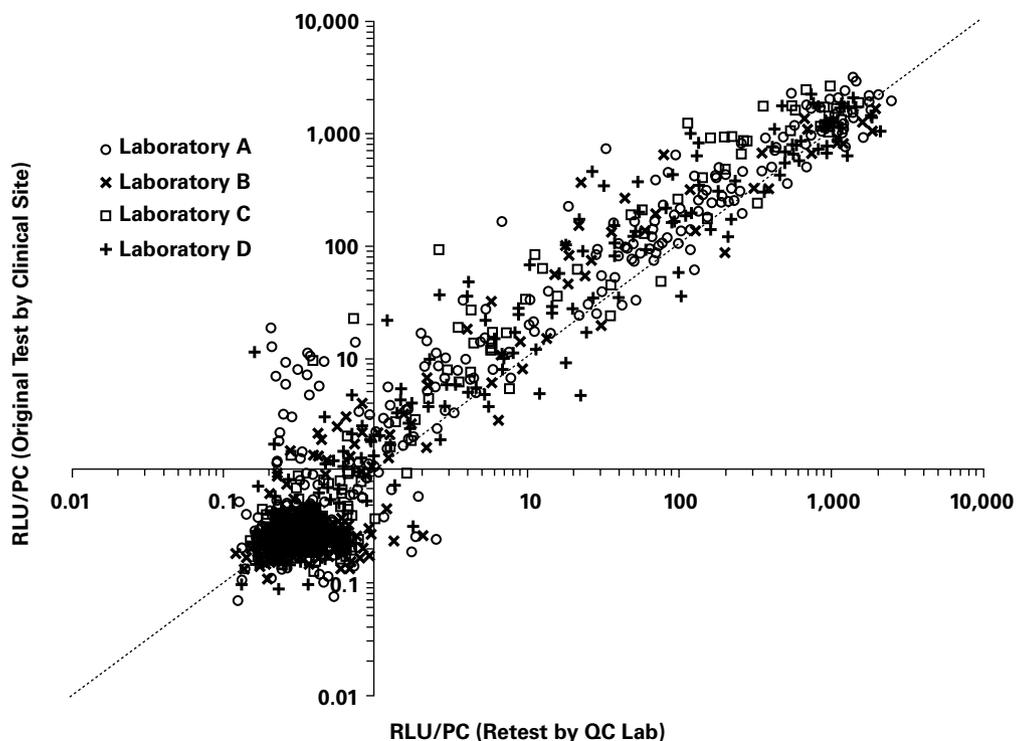
## Results

A comparison of crude RLU/PC values of the HC2 test by all 4 CC laboratories combined and of the repeated test by the HPV QC laboratory is shown in **Figure 1**. The HPV QC laboratory had lower RLU/PC values than the CC laboratories, with mean and median differences of 43.67 RLU/PC (range, 16.41-66.61 RLU/PC for individual CC laboratories) and 0.06, respectively (range, 0.02-0.29 RLU/PC for individual CC laboratories) ( $SD = 206.87$  RLU/PC), but this difference was not statistically significant ( $P = .6$ ). The overall Spearman correlation was 0.82 (95% confidence interval [CI], 0.80-0.84), and the Spearman correlation restricted to specimens interpreted as positive ( $\geq 1.0$  RLU/PC) by CC and HPV QC laboratories was 0.94 (95% CI, 0.93-0.95). Spearman correlations between individual laboratories and the HPV QC laboratory ranged from 0.73 (laboratory C) to 0.87 (laboratory A). Spearman correlations between individual laboratories and the HPV QC laboratory

restricted to specimens that twice tested positive ranged from 0.93 (laboratory D) to 0.96 (laboratory A).

There was good to excellent overall agreement of HC2 test positivity between the 4 CC laboratories as a group and the HPV QC laboratory, with a  $\kappa$  value of 0.84 (95% CI, 0.81-0.87), percentage of agreement of 92.2% (95% CI, 90.4%-93.7%), and percentage of positive agreement of 82.5% (95% CI, 78.8%-85.8%) **Table 2**. For comparison between the individual CC laboratories and the HPV QC laboratory, the  $\kappa$  values ranged from 0.78 (laboratory B) to 0.89 (laboratory C) and were inversely related to the percentage of women with negative ThinPrep cytologic results as interpreted by the CC pathologists. The test positivity for the HPV QC laboratory was less than the test positivity for the CC laboratories combined and less than for individual CC laboratories (Table 2). There were no significant differences in the reproducibility of the assay before (2-mL equivalent) and after the introduction of the smaller aliquots (1.0-mL equivalent), although there was a slight decrement in the overall signal strength (data not shown).

The lowest interlaboratory agreement was observed among results for women with negative cytologic findings, with a  $\kappa$  value of 0.73 (95% CI, 0.66-0.79) **Table 3**. The  $\kappa$



**Figure 1** Comparison of raw RLU/PC values for paired Hybrid Capture 2 test results by the clinical center laboratories and the human papillomavirus quality control (QC) laboratory ( $N = 1,072$ ). The axes are shown at 1 RLU/PC, the positive cut point for Hybrid Capture 2. Overall Spearman  $\rho = 0.82$  (95% confidence interval [CI], 0.80-0.84 [all test results]); Spearman for positive results (RLU/PC  $\geq 1.0$ ),  $\rho = 0.94$  (95% CI, 0.93-0.95);  $\kappa = 0.84$  (95% CI, 0.81-0.87). The dotted line indicates test equivalence. RLU/PC, signal strengths in relative light units compared with 1 pg/mL HPV type 16 DNA positive control samples (1.0 RLU/PC  $\sim 1$  pg/mL). See text for proprietary information.

**Table 2**  
Comparison of CC Laboratories and HPV QC Laboratory Paired Hybrid Capture 2 Test Results\*

Laboratory	No.	Cytologically Negative (%) <sup>†</sup>	Test Positivity (≥1.0 RLU/PC)			CC/QC Result				κ (95% CI)
			CC	QC	P <sup>‡</sup>	-/-	+/-	-/+	+/+	
A	328	65.5	82 (55.5)	163 (49.7)	.001	138 (42.1)	27 (8.2)	8 (2.4)	155 (47.3)	0.79 (0.72-0.85)
B	164	61.7	69 (42.1)	58 (35.4)	.008	92 (56.1)	14 (8.5)	3 (1.8)	55 (33.5)	0.78 (0.69-0.88)
C	249	53.0	78 (31.3)	71 (28.5)	.03	169 (67.9)	9 (3.6)	2 (0.8)	69 (27.7)	0.89 (0.83-0.96)
D	331	55.6	134 (40.5)	119 (36.0)	.001	194 (58.6)	18 (5.4)	3 (0.9)	116 (35.0)	0.87 (0.81-0.92)
All	1,072	59.0	463 (43.19)	411 (38.34)	<.0001	593 (55.32)	68 (6.34)	16 (1.49)	395 (36.85)	0.84 (0.81-0.87)

CC, clinical center; CI, confidence interval; HPV, human papillomavirus; QC, quality control; RLU/PC, signal strengths in relative light units compared with 1 pg/mL HPV type 16 DNA positive control samples (1.0 RLU/PC ~1 pg/mL); +, positive; -, negative.

\* Data are given as number (percentage). See text for proprietary information.

<sup>†</sup> CC pathology cytologic results were missing for 7 women (0.7%).

<sup>‡</sup> McNemar  $\chi^2$  test. All values were statistically significant ( $P < .05$ ).

values for ASC, LSIL, and HSIL or more severe categories were 0.80 or higher.

We used median RLU/PC values, PCR test results, and worst QC histologic result during the study to investigate discordant test results (CC-positive/QC-negative vs CC-negative/QC-positive), which occurred in 7.84% of the paired tests (Table 4). Comparing the median RLU/PC values for the positive tests in the discordant paired results, the CC-positive results (2.10 RLU/PC) were slightly albeit significantly greater than the QC-positive results (1.46 RLU/PC) ( $P = .02$ ). By contrast, these values for discordant tests were at least 35-fold less than the corresponding median values for each test (CC = 164.22 RLU/PC; QC = 81.05 RLU/PC) in paired tests that were twice positive ( $P = .0001$ ). Although a positive test by the CC laboratory and a negative result by the HPV QC laboratory was more common ( $n = 68$ ) than the reverse ( $n = 16$ ), there was no evidence that either positive result in a discordant pair was more likely to be associated with an oncogenic HPV type as detected by PCR ( $P = .6$ ). Of note, the oncogenic HPV positivity rate in these discordant paired results, CC-positive/QC-negative (39% [26/66]) and CC-negative/QC-positive (31%

[4/13]), was much lower than in the double-positive (84.0%) and was much greater than in the double-negative (10.3%) results ( $P < .0001$ ). Finally, 9% (5/57) of the CC-positive/QC-negative paired test results were CIN 3 or cervical cancer (vs less severe than CIN 3) compared with 7% (1/15) of the CC-negative/QC-positive results ( $P = .8$ ).

## Discussion

We describe a systematic multilaboratory study of the interlaboratory reproducibility of HC2, a clinical test being used for triage of women with equivocal Pap results and introduced for general cervical cancer screening. It is important to note that HC2 demonstrated good to excellent interlaboratory agreement, with the observed lowest  $\kappa$  value of 0.78 for one clinical site, in a group of women enrolled in the study because of their equivocal or mildly abnormal Pap smear findings. As a point of comparison, only a moderate reproducibility was reported between the CCs and the pathology QC group ( $\kappa = 0.46$ ) for evaluation of enrollment ThinPrep cytologic results for ALTS patients when distinctions of cytologic

**Table 3**  
Comparison of CC Laboratories and HPV QC Laboratory Paired Hybrid Capture 2 Test Results Including  $\kappa$  Values Stratified on CC ThinPrep Cytology Interpretation\*

CC Interpretation <sup>†</sup>	Test Positivity (≥1.0 RLU/PC)			CC/QC Result				κ (95% CI)
	CC	QC	P <sup>‡</sup>	-/-	+/-	-/+	+/+	
Negative ( $n = 628$ )	163 (26.0)	128 (20.4)	<.0001	452 (72.0)	48 (7.6)	13 (2.1)	115 (18.3)	0.73 (0.66-0.79)
ASC ( $n = 238$ )	118 (49.6)	105 (44.1)	.0008	119 (50.0)	14 (5.9)	1 (0.4)	104 (43.7)	0.87 (0.81-0.94)
LSIL ( $n = 149$ )	133 (89.3)	131 (87.9)	.4	14 (9.4)	4 (2.7)	2 (1.3)	129 (86.6)	0.80 (0.65-0.95)
HSIL or higher ( $n = 50$ )	46 (92)	45 (90)	.3	4 (8)	1 (2)	0 (0)	45 (90)	0.80 (0.64-1.00)

ASC, atypical squamous cells; CC, clinical center; CI, confidence interval; HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion; QC, quality control; RLU/PC, signal strengths in relative light units compared with 1 pg/mL HPV type 16 DNA positive control samples (1.0 RLU/PC ~1 pg/mL); +, positive; -, negative.

\* Data are given as number (percentage). See text for proprietary information.

<sup>†</sup> Cytology results were missing for 7 women (0.7%).

<sup>‡</sup> McNemar  $\chi^2$  test. Statistical significance was set at  $P < .05$ .

**Table 4**  
**Comparison of Paired HC2 Test Results With Median RLU/PC values, PGMY09/11 PCR Results, and Worst QC Histologic Diagnosis During the Entire ALTS\***

	CC/QC Test Result				Total (N = 1,072)
	-/- (n = 593)	+/- (n = 68)	-/+ (n = 16)	+/+ (n = 395)	
Median RLU/PC value					
CC	0.28	2.10	0.38	164.22	—
QC	0.31	0.48	1.46	81.05	—
PCR results (hierarchical) <sup>†</sup>					
Oncogenic	41 (10.3)	26 (39)	4 (31)	315 (84.0)	386
Nononcogenic	97 (24.3)	17 (26)	5 (38)	40 (10.7)	159
Negative	261 (65.4)	23 (35)	4 (31)	20 (5.3)	308
Total	399	66	13	375	853
Worst QC histologic diagnosis					
CIN 3 or worse	4 (0.8)	5 (9)	1 (7)	57 (15.3)	67
Less severe than CIN 3	519 (99.2)	52 (91)	14 (93)	315 (84.7)	900
Total	523	57	15	372	967

ALTS, ASCUS (atypical squamous cells of undetermined significance) and LSIL (low-grade squamous intraepithelial lesion) Triage Study; CIN, cervical intraepithelial neoplasia; CC, clinical center; HC2, Hybrid Capture 2; HPV, human papillomavirus; PCR, polymerase chain reaction; PGMY09/11, L1 consensus primer; QC, quality control; RLU/PC, signal strengths in relative light units compared with 1 pg/mL HPV type 16 DNA positive control samples (1.0 RLU/PC ~ 1 pg/mL); +, positive; -, negative.

\* Data are given as number (percentage) unless otherwise indicated. See text for proprietary information. *P* values for discordant results (+/- vs -/+) were as follows: median RLU/PC value, *P* = .02 (Kruskal-Wallis); PCR results, *P* = .6 (Pearson  $\chi^2$ ); worst QC histologic diagnosis, *P* = .8 (Pearson  $\chi^2$ ).

<sup>†</sup> PCR results were ranked hierarchically according to cancer risk: oncogenic (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68) > nononcogenic > negative. Specimens with both oncogenic and nononcogenic types were classified as oncogenic.

severity were made,<sup>14</sup> and the reproducibility was only slightly improved when cytologic findings were categorized as negative or abnormal (ASC or more severe) ( $\kappa = 0.56$ ).

There are surprisingly few data on the reproducibility of HC2 despite its introduction as an adjunct to cytologic examination for cervical cancer screening in women older than 30 years.<sup>2,5,15</sup> One intralaboratory evaluation reported a  $\kappa$  value of 0.72 when using a nonstandard specimen, a cervicovaginal lavage, in the context of a screening population.<sup>16</sup> A recent study retested 159 of 649 positive HC2 specimens (24.5%) and found that 70% retested negative ( $\geq 17\%$  of all HC2-positive specimens). It is noteworthy that the majority of retest negative results (69%) had an initial RLU/PC value in the 1 to 3 range.<sup>17</sup> However, this study did not systematically evaluate reproducibility (eg, retest negative specimens). In our study, we found that 68 (14.7%) of 463 positive specimens had negative results when retested, and 44 (65%) of the 68 had an original RLU/PC value between 1 and 3.

We cannot explain the systematically lower RLU/PC on retesting by the HPV QC laboratory. We note that the reproducibility did not differ significantly by test batch (data not shown) or by the CC that obtained the specimen, suggesting consistent test performance by the HPV QC laboratory over time and testing independent of specimen origin. Although signal strength was affected by switching to a smaller specimen equivalent after September 1998, there were no observable differences in test agreement between CC and QC laboratories before and after the change, even with stratification on cytologic interpretation (data not shown), ruling out this change as a main effect.

As the result of the study design in which the HPV QC laboratory always tested specimens after testing by the CC laboratories, we suggest that it is the sequence of testing that affected the performance. One possible explanation is specimen degradation between test and retest. Although it previously has been shown that there is some degradation of these specimens over many years,<sup>18,19</sup> the interval between tests was much shorter (mean, 179 days; median, 181 days; range, 59-379 days) than the timeframe (several years) of the observed degradation. In the present study, we observed a weak, nonsignificant association (*P* = .2) of the intervening time between CC and QC testing and the difference in signal strength (RLU/PC) between the 2 tests, which might partially explain these differences in test performance (data not shown). We note that HC2 testing for HPV DNA is recommended only on residual PreservCyt specimens less than 30 days old, a shorter interval than the shortest interval between the test and retest (59 days). Thus, it is conceivable that signal degradation occurred in the interval before retesting and we were unable to observe it in this study.

Another possible explanation is that PreservCyt specimens are nonhomogeneous, and, as a result, there are lower concentrations of HPV-infected cells (per unit volume) as the specimen is used up, thereby reducing HC2 positivity. Such an effect, not yet documented, likely would be dependent on the method of specimen preparation and processing.

Regardless of the cause, the decrease in signal strength on retesting has little clinical significance because HPV retesting is not part of routine clinical practice. However, some laboratories have adopted a policy of retesting "low"

positive results ( $\leq 5$  RLU/PC), which warrants caution given the greater likelihood of discordant test results. Given the unreliability of retesting specimens with low positive results and the uncertain interpretation of discordant test results, it might be safer to refer women with ASC and a low positive test result by HC2 to colposcopy rather than perform a potentially uninformative and expensive HC2 retest.

One limitation of this study was that we did not assess reproducibility in the specimens from the general population, only from women referred into the study with equivocal or mildly abnormal cytologic findings who might be expected on average to have higher viral loads than those with predominantly normal cytologic findings.<sup>1</sup> As expected with any assay with a positive cut point, the signal-to-noise ratio is lowest as the signal strength approaches this cut point. As a case in point, the lowest reproducibility ( $\kappa = 0.73$ ) was observed in the results for women whose samples were judged to be cytologically negative by CC pathologists', similar to the aforementioned intralaboratory results in a predominantly negative screening population.<sup>16</sup> Thus, we anticipate slightly poorer reliability in general screening, primarily owing to a greater number of specimens with lower viral loads and concomitant signal strengths near the positive threshold. It is important to note that the reliability is excellent among results for women with equivocal or abnormal cytologic findings who have a greater likelihood of having underlying CIN 3 or more severe disease than women with negative cytologic findings.

We used raw HC2 values, PCR data, and histologic diagnoses to assist in our interpretation of the discordant test results. Not surprisingly, given the lower signal on the retest, there were more CC-positive/QC-negative paired tests than the converse, and CC-positive/QC-negative paired test results were slightly more likely to be positive for oncogenic HPV DNA as detected by PCR. As alluded to in the preceding discussion, both pairs of discordant test results were characterized by a lower signal strength, a surrogate of viral load,<sup>20</sup> in the positive test. Likewise, of the cases of CIN 3 or cervical cancer that were positive by one or the other test but not both, more than 90% had RLU/PC values less than 10.0. Discordant test results also were characterized by lower PCR positivity for oncogenic types and a high PCR positivity for nononcogenic types compared with specimens with positive results on both tests. It is possible that these sporadic positive test results were false-positive results due to cross-reactivity with nononcogenic types.<sup>21</sup> However, we did not find any single nononcogenic HPV type associated with the discordant HC2 test results.

We demonstrated that HC2 is a reliable clinical test for HPV DNA detection despite a slight systematic decrement in the retest signal that undoubtedly reduced the agreement between these 2 tests.

*From the Divisions of<sup>1</sup>Cancer Epidemiology and Genetics and<sup>3</sup>Cancer Prevention, National Cancer Institute, National Institutes of Health, DHHS, Bethesda, MD; and<sup>2</sup>Departments of Molecular Genetics and Microbiology and Obstetrics and Gynecology, University of New Mexico Health Sciences Center, School of Medicine, Albuquerque.*

*Supported by contracts CN-55153, CN-55154, CN-55155, CN-55156, CN-55157, CN-55158, CN-55159, and CN-55105 with the National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD. Some of the equipment and supplies used in the study were donated or provided at reduced cost by Digene, Gaithersburg, MD; Cytoc, Boxborough, MA; National Testing Laboratories, Fenton, MO; Denvu, Tucson, AZ; TriPath Imaging, Burlington, NC; and Roche Molecular Systems, Alameda, CA.*

*Address reprint requests to Dr Wheeler: University of New Mexico, Molecular Genetics and Microbiology, Room 347, 915 Camino de Salud NE, Albuquerque, NM 87131-5276.*

*\* National Cancer Institute, Bethesda, MD: D. Solomon, project officer; M. Schiffman, coproject officer; R. Tarone, statistician; Clinical Centers: University of Alabama at Birmingham: E.E. Partridge, principal investigator; L. Kilgore, co-principal investigator; S. Hester, study manager; University of Oklahoma, Oklahoma City: J.L. Walker, principal investigator; G.A. Johnson, co-principal investigator; A. Yadack, study manager; Magee-Women's Hospital of the University of Pittsburgh Medical Center Health System, Pittsburgh, PA: R.S. Guido, principal investigator; K. McIntyre-Seltman, co-principal investigator; R.P. Edwards, investigator; J. Gruss, study manager; University of Washington, Seattle: N.B. Kiviat, co-principal investigator; L. Koutsky, co-principal investigator; C. Mao, investigator; J.M. Haug, study manager; Colposcopy Quality Control Group: D. Ferris, principal investigator, Medical College of Georgia, Augusta; J.T. Cox, coinvestigator, University of California at Santa Barbara; L. Burke, coinvestigator, Beth Israel Deaconess Medical Center Hospital, Boston, MA; HPV [human papillomavirus] Quality Control Group: C.M. Wheeler, principal investigator, and C. Peyton-Goodall, laboratory manager, University of New Mexico Health Sciences Center, Albuquerque; M.M. Manos, coinvestigator, Kaiser Permanente, Oakland, CA; Pathology Quality Control Group: R.J. Kurman, principal investigator, and D.L. Rosenthal, coinvestigator, Johns Hopkins Hospital, Baltimore, MD; M.E. Sherman, coinvestigator, National Cancer Institute, Rockville, MD; M.H. Stoler, coinvestigator, University of Virginia Health Science Center, Charlottesville; Westat, Coordinating Unit, Rockville, MD: J. Rosenthal, project director; M. Dunn, data management team leader; J. Quarantillo, senior systems analyst; D. Robinson, clinical center coordinator; Quality of Life Group: D.M. Harper, chair, Dartmouth Medical School, Hanover, NH; Digene, Gaithersburg, MD: A.T. Lorincz, senior scientific officer; Information Management Services, Silver Spring, MD: B. Kramer, senior programmer/analyst.*

## References

1. Schiffman M, Herrero R, Hildesheim A, et al. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. *JAMA*. 2000;283:87-93.

2. Cuzick J, Szarewski A, Cubie H, et al. Management of women who test positive for high-risk types of human papillomavirus: the HART study. *Lancet*. 2003;362:1871-1876.
3. Solomon D, Schiffman M, Tarone R. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. *J Natl Cancer Inst*. 2001;93:293-299.
4. ASCUS-LSIL Triage Study (ALTS) Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. *Am J Obstet Gynecol*. 2003;188:1383-1392.
5. Saslow D, Runowicz CD, Solomon D, et al. American Cancer Society guideline for the early detection of cervical neoplasia and cancer. *CA Cancer J Clin*. 2002;52:342-362.
6. Schiffman M, Adriaan ME. ASCUS-LSIL Triage Study: design, methods and characteristics of trial participants. *Acta Cytol*. 2000;44:726-742.
7. ASCUS-LSIL Triage Study (ALTS) Group. A randomized trial on the management of low-grade squamous intraepithelial lesion cytology interpretations. *Am J Obstet Gynecol*. 2003;188:1393-1400.
8. Gravitt PE, Peyton CL, Alessi TQ, et al. Improved amplification of genital human papillomaviruses. *J Clin Microbiol*. 2000;38:357-361.
9. Peyton CL, Gravitt PE, Hunt WC, et al. Determinants of genital human papillomavirus detection in a US population. *J Infect Dis*. 2001;183:1554-1564.
10. Bosch FX, Manos MM, Munoz N, et al, for the International Biological Study on Cervical Cancer (IBSCC) Study Group. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J Natl Cancer Inst*. 1995;87:796-802.
11. Cox JT, Schiffman M, Solomon D. Prospective follow-up suggests similar risk of subsequent cervical intraepithelial neoplasia grade 2 or 3 among women with cervical intraepithelial neoplasia grade 1 or negative colposcopy and directed biopsy. *Am J Obstet Gynecol*. 2003;188:1406-1412.
12. Guido R, Schiffman M, Solomon D, et al. Postcolposcopy management strategies for women referred with low-grade squamous intraepithelial lesions or human papillomavirus DNA-positive atypical squamous cells of undetermined significance: a two-year prospective study. *Am J Obstet Gynecol*. 2003;188:1401-1405.
13. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA*. 2002;287:2114-2119.
14. Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA*. 2001;285:1500-1505.
15. Ressel GW, American College of Obstetricians and Gynecologists. ACOG releases guidelines on cervical cytology screening. *Am Fam Physician*. 2003;68:2081, 2084.
16. Castle PE, Lorincz AT, Mielzynska-Lohnas I, et al. Results of human papillomavirus DNA testing with the Hybrid Capture 2 assay are reproducible. *J Clin Microbiol*. 2002;40:1088-1090.
17. de Cremoux P, Coste J, Sastre-Garau X, et al. Efficiency of the hybrid capture 2 HPV DNA test in cervical cancer screening: a study by the French Society of Clinical Cytology. *Am J Clin Pathol*. 2003;120:492-499.
18. Castle PE, Solomon D, Hildesheim A, et al. Stability of archived liquid-based cervical cytologic specimens. *Cancer*. 2003;99:89-96.
19. Castle PE, Hildesheim A, Schiffman M, et al. Stability of archived liquid-based cytologic specimens [letter]. *Cancer*. 2003;99:320-322.
20. Gravitt PE, Burk RD, Lorincz A, et al. A comparison between real-time polymerase chain reaction and Hybrid Capture 2 for human papillomavirus DNA quantitation. *Cancer Epidemiol Biomarkers Prev*. 2003;12:477-484.
21. Castle PE, Schiffman M, Burk RD, et al. Restricted cross-reactivity of Hybrid Capture 2 with nononcogenic human papillomavirus types. *Cancer Epidemiol Biomarkers Prev*. 2002;11:1394-1399.