

Evaluation of Self-Collected Cervicovaginal Cell Samples for Human Papillomavirus Testing by Polymerase Chain Reaction

Patti E. Gravitt,¹ James V. Lacey, Jr., Louise A. Brinton, Willard A. Barnes, Janet R. Kornegay, Mitchell D. Greenberg, Sarah M. Greene, Olympia C. Hadjimichael, Larry McGowan, Rodrigue Mortel, Peter E. Schwartz, Richard Zaino, and Allan Hildesheim

Department of Human Genetics, Roche Molecular Systems, Alameda, California 94501 [P. E. G., J. R. K.]; Environmental Epidemiology Branch, National Cancer Institute, Rockville, Maryland 20852-7234 [P. E. G., J. V. L., L. A. B., A. H.]; Lombardi Cancer Center, Georgetown University, Washington, DC 20057 [W. A. B.]; Graduate Hospital, Philadelphia, Pennsylvania 19146 [M. D. G.]; Westat, Inc., Rockville, Maryland 20850 [S. M. G.]; Yale University School of Medicine, New Haven, Connecticut 06520 [O. C. H., P. E. S.]; George Washington University, Washington, DC 20052 [L. M.]; and Milton S. Hershey Medical Center, Hershey, Pennsylvania 17033 [R. M., R. Z.]

Abstract

As human papillomavirus (HPV) becomes accepted as the central cause of cervical cancer, longitudinal studies are shifting focus away from causality to a more detailed investigation of the natural history of HPV infections. These studies commonly require repeated samples for HPV testing over several years, usually collected during a pelvic exam, which is inconvenient to the participants and costly to the study. To alleviate the inconvenience and cost of repeated clinic visits, it has been proposed that women collect cervicovaginal cells themselves, hopefully increasing participation in the natural history studies. We evaluated the technical feasibility of self-collection of cervicovaginal cells using a Dacron swab for HPV DNA detection. We compared the self-collected swab sample and two clinician-administered swab samples (one from the endocervix and another from the ectocervix) from a total of 268 women participating in a case-control study of adenocarcinoma and squamous cell carcinomas of the uterine cervix (111 cases and 157 controls). HPV DNA was detected and genotyped using an L1 consensus PCR assay. The overall agreement between the clinician- and self-collected swabs was excellent [88.1%; $\kappa = 0.73$ (95% confidence interval (CI), 0.61–0.85)]. The correlation was highest between the two clinician-administered swabs [$\kappa = 0.81$ (95% CI, 0.69–0.93)] but was still excellent when comparing either clinician-administered swab to the self-administered sample [$\kappa = 0.75$ (95% CI, 0.63–0.87) and 0.67 (95% CI, 0.55–0.79) for ectocervix and

endocervix, respectively]. The type-specific agreement between samples was higher for high-risk, or cancer-associated, HPV genotypes than for low risk, noncancer-associated HPV genotypes when comparing the self-administered swab sample to the clinician-administered swab sample ($\kappa = 0.78$ for high-risk versus 0.66 for low-risk HPV infections, $t = -1.45$, $P = 0.15$). The decrease in agreement for low risk types was largely attributable to an increased detection of these types in the self-administered sample (McNemar's $\chi^2 = 6.25$, $P = 0.01$ for clinician- versus self-administered swab comparisons). The agreement did not vary significantly by age, menopausal status, case status, or clinic center. We have demonstrated that a self-collected Dacron swab sample of cervicovaginal cells is a technically feasible alternative to clinician-administered cervical cell collection in natural history studies of HPV and cervical cancer.

Introduction

Case-control and cohort epidemiological study designs were instrumental in defining HPV² as the central cause of cervical cancer (1–3). Similar study designs are now being used in more detailed investigations of the natural history of HPV infections (4–6). A common problem is the ability to recruit women to participate in these studies, particularly those that are longitudinal in design, requiring multiple follow-up contacts with each participant. Low response rates among women without cervical disease (controls) have historically been attributable to the reluctance to volunteer for routine pelvic examinations that are required to collect cervicovaginal cells for HPV testing. To encourage participation and compliance with follow-up, some investigators have begun to assess the feasibility of having the participant collect her own cervical cells that can then be sent to the laboratory for HPV DNA testing, circumventing the clinic visit and pelvic examination altogether. A variety of collection instruments have been evaluated, including Dacron swabs (7), tampons (8–10), and cervicovaginal lavage kits (11), with HPV DNA detection by a variety of amplified and nonamplified assays. We evaluated the correlation of HPV DNA detection from self-collected swab samples and both ecto- and endocervical directed, clinician-administered swabs using the MY09/MY11/HMBB01 L1 PCR assay (12).

Materials and Methods

Study Population. The women participating in this study were recruited as part of a matched, multicenter, case-control study of adenocarcinoma and SCC of the uterine cervix conducted in the Eastern United States. The details of study enrollment have been described (13). Briefly, SCC patients were 1:1 matched on clinic, age in 5-year intervals at diagnosis, and stage of disease at diag-

Received 12/28/99; revised 10/12/00; accepted 11/13/00.

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 To whom requests for reprints should be addressed, at Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, MSC 7234, Rockville, MD 20852-7234. E-mail: gravittp@mail.nih.gov.

² The abbreviations used are: HPV, human papillomavirus; SCC, squamous cell carcinoma; STM, standard transport medium; CI, confidence interval.

nosis to adenocarcinoma cases. Population controls were selected by random digit dialing and were individually matched to adenocarcinoma cases on age in 5-year intervals, clinic, race/ethnicity, and telephone exchange. Control women did not have cervical cancer but could have had less severe cytological abnormalities, including cervical intraepithelial neoplasia. The median age of the combined population of cases and controls was 38 years. The population was largely white (83%), with 89% of participants completing a high school education and 59% reporting some education beyond high school. A total of 203 women with adenocarcinoma (including 49 *in situ* and 142 invasive cancers, 12 with unknown disease stage) and 255 women with SCC (including 64 *in situ* and 167 invasive cancers, 24 with unknown disease stage) were identified as eligible to participate in the study. Included in the adenocarcinoma case definition were "true" adenocarcinomas, adenosquamous carcinomas, and other cervical tumors of glandular origin. Of the eligible cases, 145 of 203 (71%) of the adenocarcinoma cases and 145 of 255 (57%) of the SCC cases were enrolled. Reasons for not enrolling included refusal, too ill, not located, death, or did not speak English (13). In addition, 307 population controls were studied. All participating women signed informed consent that was approved by participating institutional review boards.

Clinician-administered Swab Collection. Each woman received a standard pelvic examination to collect cells for HPV testing. Two cervical samples were collected by the administering clinician, the first taken from the ectocervix and the second from the endocervix, using a Dacron swab. Separate endocervical and ectocervical swabs were obtained to evaluate whether endocervical infections were more likely than ectocervical infections to be missed by self-collection methods. The clinician directed the placement of the swab to the appropriate site and turned the swab two full rotations to maximize cell collection. The swab was placed into a 5-ml vial containing 1 ml of STM (Digene Diagnostics, Silver Spring, MD). For cases who had treatment by removal of the cervix prior to the HPV collection, the clinician collected two specimens from the vaginal cuff. All swabs in STM were frozen at the clinic and shipped on dry ice to the biorepository, where they were stored at -70°C .

Self-administered Swab Collection. All participating women were asked to collect a sample of cervicovaginal cells by themselves, either at home or at the clinical center. If self-collection was performed at the clinic, the study coordinators encouraged self-collection prior to the clinician-administered HPV collection and pelvic examination. Explanation and instructions for self-collection of cervicovaginal cells were given by the study personnel. In brief, each woman was given a long-handled, sterile Dacron swab in a wrapper (the same type of swab used by the clinician), a plastic 50-ml vial with a lid, and a paper towel with a plastic lining on one side. Subjects were directed to a private, well-lit room. Verbal and written instructions for cell collection, as well as a diagram of the female genital anatomy, were provided. "Remove clothes from the waist down, remove the swab from the wrapping, and choose a comfortable position (either standing with one foot on the toilet or bathtub, or standing with legs apart and knees slightly bent). Relax and insert the cotton tip of the swab into the vagina, without touching the labia or urethra if possible. Gently push the swab up into the vagina until physically it cannot go any further. Using your thumb and two fingers, pull the swab halfway out of the vagina, and then re-insert it. Rotate the swab inside the vagina for three full rotations, keeping the swab as far into the vagina as possible. Withdraw the swab holding the lips of the labia apart and taking care not to touch other portions of the genitals, and place the swab directly into the plastic vial that is provided." If any irritation or

discomfort was encountered, the women were instructed to reduce the pressure of the swab inside the vagina, pull the swab out away from the cervix a bit, or stop the procedure completely. After the subject had completed the sample collection, the study personnel removed the swab from the 50-ml plastic vial and transferred it carefully to the 5-ml vial containing 1 ml of STM. Once transferred, the samples were in vials and transport medium identical to those collected by the clinicians. The swabs were stored at -70°C until processing.

Exclusions. All participating women were asked to contribute three samples for HPV testing: one swab taken from the endocervix and one swab taken from the ectocervix, both collected by the clinician while administering a standard pelvic examination, and one self-administered swab taken by the woman herself as instructed by study personnel. Of the 145 women with adenocarcinoma of the cervix who were enrolled in the study, 138 (95%) contributed at least one of the three samples for HPV testing. Of these 138 women, 17 (12%) did not contribute a self-administered swab sample, and 13 (9%) did not contribute a clinician-administered sample, resulting in a total of 108 adenocarcinoma cases (53% of total eligible; 74% of total enrolled) with all three samples for comparison. Of the 145 women with SCC enrolled in the study, 136 (94%) contributed at least one sample for HPV testing. Fifteen women (11%) did not contribute a clinician-administered swab, and 12 women (9%) did not contribute a self-administered swab, leaving 109 total SCC cases with all three samples (43% of total eligible; 75% of total enrolled). Of the 307 control women, 255 (83%) contributed at least one sample for HPV testing. Of these 255 controls, 98 (38%) submitted only the self-collected swab sample, leaving a total of 157 controls with all three swab samples (51% of total eligible). The women excluded because of missing swab samples were similar to the women remaining in the analysis with regard to age, level of education, and race (data not shown). The women with missing samples were more likely to be controls, with a lower HPV prevalence (27% HPV positive among those excluded *versus* 39% HPV positive among those remaining in the analysis; $P = 0.008$).

We further restricted the analysis to women whose clinician-administered swabs were collected from an intact cervix (*i.e.*, women who were enrolled into the case-control study prior to surgical treatment for cancer). The total number of adenocarcinoma cases was reduced to 45 women, after exclusion of 58 women with only vaginal samples and 5 women with unknown sample type. The total number of SCC cases was reduced to 66, after exclusion of 42 women with vaginal clinician-administered samples and 1 woman with unknown sample type. All control women with three swab samples had clinician-administered swabs directed to the cervix. The 100 women who were excluded because of vaginal clinician-administered swabs were exclusively cases (58 adenocarcinomas and 42 SCCs) and were older than the women remaining in the analysis (mean age, 42.2 *versus* 37.6 years; $P < 0.001$). They did not differ from the women remaining in the final analysis with respect to level of education, race, or HPV prevalence.

Sample Preparation. The cell samples in STM were processed for PCR by addition of one-half volume of Digene sample denaturation reagent (Digene Diagnostics) and incubation at 65°C for 1 h. Samples were vigorously vortexed every 15 min during the incubation. After the incubation, the tubes of the STM vial were opened carefully using a 2×2 -inch gauze pad to reduce aerosols, the swab was removed and discarded, and $150 \mu\text{l}$ were removed and aliquoted to a 1.5-ml microcentrifuge tube containing $600 \mu\text{l}$ of absolute ethanol and 1:10 volume, 5 M ammonium acetate.

Samples were mixed by inversion, and the DNA was precipitated at -20°C overnight. The precipitated DNA was pelleted by centrifugation at $12,000 \times g$ in a microcentrifuge for 30 min, and the ethanol was removed using a fine-tip plastic transfer pipette. The pellets were dried overnight and resuspended in $75 \mu\text{l}$ of Tris-EDTA buffer. Five μl of the DNA preparation were used for each amplification reaction.

HPV Testing. The presence of HPV DNA in the cell samples was determined using L1 consensus PCR, with genotype determination by a reverse line blot hybridization method (12, 14). The samples were coamplified with the biotinylated HPV L1 consensus primers, MY09/MY11/HMB01, and β -globin primers, GH20 and PC04. Amplification was performed as described previously (12) in master mix containing $1 \times$ PCR Buffer II, 6 mM MgCl_2 , $200 \mu\text{M}$ each dATP, dGTP, and dCTP, $600 \mu\text{M}$ dUTP, 7.5 units of AmpliTaq Gold (Perkin-Elmer, Foster City, CA), 50 pmol each of MY09 and MY11, and 5 pmol each of HMB01, GH20 and PC04. Reactions were performed in a GeneAmp PCR System 9600 thermal cycler (Perkin-Elmer, Foster City, CA) using the following temperature profile: 95°C AmpliTaq Gold activation for 9 min; 40 cycles of 95°C for 1 min, 55°C for 1 min, 72°C for 1 min; 72°C final extension for 5 min, and stored at $4-15^{\circ}\text{C}$ until the detection procedure.

PCR products were denatured by addition of equal volume 0.4 N NaOH. Seventy-five μl of denatured amplicon were added to each genotyping strip. Hybridization was performed in $4 \times$ SSPE/ 0.1% SDS at 53°C for 30 min, followed by a stringent wash in $1 \times$ SSPE/ 0.1% SDS at 53°C for 15 min. Positive hybridization of the biotinylated PCR product was detected by blue color deposition at the probe site after streptavidin-horseradish peroxidase conjugation and color development as described previously (12). Samples were considered sufficient for HPV analysis if they had a positive β -globin hybridization. Samples were considered to be HPV positive if they hybridized with one or more of the 27 type-specific HPV probes on the genotyping strip. Samples that did not hybridize to any of the 27 HPV probes were considered HPV negative for this analysis. We did not extend the analysis to include the detection of other genital HPV types, because such types are not known to be associated with a significant fraction of cervical cancer (15).

HPV 16 Contamination Resolution. The original amplification and hybridization results indicated a disproportionately high prevalence of HPV 16 among the control women and a lack of association between HPV presence and sexual behavior. On the basis of results from an earlier study that showed contamination of a particular STM lot with HPV 16 (unpublished data), we reamplified all samples that were originally positive for HPV 16 by the standard L1 consensus primer assay using a set of primers that flank the *Bam*HI cloning site in the native-type HPV 16 sequence. Using this system, wild-type HPV 16 DNA will amplify an ~ 150 -bp fragment, whereas cloned HPV 16 DNA will fail to amplify because of the insertion of a 3–4 kb vector between the primer binding sites (16). Results of this analysis confirmed the presence of HPV 16 in 18 of 88 (20%) samples originally positive for HPV 16 from STM lot 0001TX95 and in 132 of 165 (80%) samples from the remaining 4 known and all unknown STM lots (percentage confirmed by STM lots, excluding 0001TX95, ranged from 72 to 100%). In the final analysis, we considered only samples that were positive by the HPV 16 anticontamination primers to be HPV 16 positive.

Statistical Analyses. The correlation of HPV detection by the three sample collection methods was determined using an unweighted κ statistic to determine the percentage of correlation

Table 1 Agreement of HPV results between self-collected and clinician-collected swabs^a

| | Clinician-collected swab | | Total |
|--------------------------------|--------------------------|--------------|-------|
| | HPV positive | HPV negative | |
| Self-administered HPV positive | 74 | 18 | 92 |
| Self-administered HPV negative | 14 | 162 | 176 |
| Total | 88 | 180 | 268 |

^a Percentage of agreement, 236 of 268 (88.1%); percentage of agreement among positives, 74 of 106 (69.8%); $\kappa = 0.73$ (95% CI, 0.61–0.85).

beyond that expected by chance (17). The ectocervical and endocervical swab HPV results were compared separately and as a combined clinician-administered swab result to the self-administered swab HPV results. The combined result was defined as HPV negative when both ectocervical and endocervical swabs were HPV negative and HPV positive when either swab was HPV positive. When there was no difference in agreement with the self-administered swab in the ectocervical and endocervical strata, only the combined clinician-administered swab results are presented. A *t* statistic was calculated using the SE for κ (17) to test differences in κ estimates. A McNemar's χ^2 test for matched pairs was used to test the significance of unequal distribution of discordant results (18). Pearson χ^2 statistics were used to compare the distribution of exposures between groups (19). Exact tests for symmetry from $K \times K$ tables were used to test for differences in total number of HPV genotypes detected per sample between the different swab collection methods (20). All data analysis procedures were computed using the Stata 6.0 software package (21).

Results

The overall prevalence of HPV (defined as positive in at least one sample) in the final population of 268 women was 39.6%. As expected, HPV prevalence was highest in women with cancer (62% among adenocarcinoma cases and 71% among SCC cases) relative to control women (20%). The modest HPV prevalence among cases in this study likely reflects the collection of samples for HPV testing from some women after treatment of their cervical cancer. Only one sample from each of the three collection methods was insufficient for HPV DNA analysis because of lack of β -globin amplification. Each of the three β -globin-negative results was from a different subject. The correlation of HPV from clinician-collected versus self-collected swabs is summarized in Table 1. The crude HPV agreement was 88.1% total (236 of 268) and 69.8% (74 of 106) among the HPV-positive results ($\kappa = 0.73$; 95% CI, 0.61–0.85). Of the 106 women with at least one positive HPV test result, 74 of 106 (70%) were detected by both collection methods. Fourteen of the 106 total positive women (13%) were “missed” by self-collected swabs, and 18 of 106 (17%) were “missed” by clinician-collected swabs (McNemar's $\chi^2 = 0.50$; $P = 0.48$). When restricting the analysis to pairwise comparisons, the agreement was highest when comparing the two clinician-administered swabs (*i.e.*, the endocervical versus ectocervical swabs) with 92.2% overall agreement and 76.4% agreement among the positives ($\kappa = 0.81$; 95% CI, 0.69–0.93). The discordant results were equally distributed between the ectocervix and endocervix (McNemar's $\chi^2 = 0.43$; $P = 0.51$). The ectocervical samples were marginally better correlated with the self-administered samples than the endocervical swabs ($\kappa = 0.75$ versus 0.67; $t = 0.93$; $P = 0.35$). Among the discordant results, the self-administered swabs were more likely to be positive than either the ectocervical or the endocervical swabs (McNemar's $\chi^2 = 5.83$ and

Table 2 Type-specific agreement between clinician- and self-administered swab collection for HPV types with five or more positive results

| HPV type | % of agreement ^a | κ | 95% CI |
|----------------------------|-----------------------------|----------|-----------|
| Any low risk ^b | 18/34 (52.9%) | 0.66 | 0.54–0.78 |
| 6 | 2/6 (33.3%) | 0.49 | 0.39–0.60 |
| 53 | 8/13 (61.5%) | 0.75 | 0.64–0.87 |
| 54 | 7/13 (53.8%) | 0.69 | 0.57–0.81 |
| Any high risk ^c | 67/91 (73.6%) | 0.78 | 0.67–0.90 |
| 16 | 38/55 (69.1%) | 0.78 | 0.66–0.90 |
| 18 | 11/15 (73.3%) | 0.84 | 0.72–0.96 |
| 55 | 4/5 (80.0%) | 0.89 | 0.77–1.01 |
| 58 | 5/6 (83.3%) | 0.91 | 0.79–1.03 |
| MM7 | 4/10 (40.0%) | 0.56 | 0.44–0.68 |

^a Among the positives.

^b HPV 6, 11, 40, 42, 53, 54, 57, 66, or MM8.

^c HPV 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 55, 56, 58, 59, 68, MM4, MM7, or MM9.

6.74 for ectocervix and endocervix, respectively; $P < 0.05$ for both).

The overall HPV typing agreement between the self-administered and the clinician-administered swabs was good, with perfect type agreement in 54 of 106 (51%) positive samples and partial type agreement (at least one type detected in both samples) in 19 of 106 positive samples (19%). There was complete discordance in 33 of 106 (31%) of the positive samples, but 32 of 33 of these discordant types represented the discordance seen in HPV positivity in Table 1. Of the positive samples by each collection method, 27 of 92 (29.3%) of the self-administered and 26 of 88 (29.5%) of the clinician-administered HPV-positive women were infected with more than one HPV genotype. We compared the total number of types detected per sample (range, 1–7) in women with at least one positive HPV test using an exact test for symmetry. The results of this analysis showed a marginal increase in multiple types per sample with the self-administered *versus* either clinician-administered swab sample ($P = 0.07$ and 0.03 for ectocervical and endocervical swabs, respectively). The majority of the asymmetry was attributable to the discordant pairs with zero (HPV negative) *versus* one type, indicating no particular pre-dilection of any swab type to detect multiple HPV genotypes.

We examined the agreement by individual HPV genotype for HPV types that had at least five positive results: HPV 6, 16, 18, 53, 54, 55, 58, and MM7 (Table 2). The agreement ranged from moderate to excellent ($\kappa = 0.49$ – 0.91). Because the types that had lower agreement between the clinician-administered swab and the self-administered swab tended to be in the non-cancer-associated or low-risk category, the type results were examined by risk group. We collapsed the HPV type-specific data to “HPV negative”; “low risk positive” if positive by any of the following low risk HPV types (6, 11, 40, 42, 53, 54, 57, 66, or MM8), regardless of coinfection by high-risk types; and “high risk positive” if positive by any of the following high-risk HPV types (16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 55, 56, 58, 59, 68, MM4, MM7, or MM9), regardless of coinfection with low-risk types. When the two clinician-administered swabs were compared by risk group, there was no difference in correlation between detectability of any low-risk *versus* any high-risk HPV type ($\kappa = 0.75$ and 0.79 for low and high risk, respectively). However, the agreement for any high-risk HPV type was much better than the agreement for low-risk types when comparing the self-administered swab *versus* either clinician-administered swab result [ectocervix *versus* self: $\kappa = 0.81$ and 0.59 for high and low risk, respectively ($t = -2.61$;

Table 3 Overall HPV agreement between clinician- and self-administered swab collection, stratified by age in 10-year categories

| Age category | <i>n</i> | % of agreement ^a | κ | 95% CI |
|--------------|----------|-----------------------------|----------|------------|
| <30 years | 68 | 22/35 (62.9%) | 0.61 | 0.37–0.85 |
| 30–39 | 92 | 21/31 (67.7%) | 0.73 | 0.53–0.94 |
| 40–49 | 66 | 18/23 (78.3%) | 0.82 | 0.59–1.06 |
| 50–59 | 27 | 7/9 (77.8%) | 0.82 | 0.45–1.19 |
| 60+ | 10 | 3/5 (60.0%) | 0.58 | –0.04–1.20 |

^a Among the positives.

$P = 0.009$); endocervix *versus* self: $\kappa = 0.72$ and 0.59 for high and low risk, respectively ($t = -1.54$; $P = 0.12$)]. This difference remained, even when the high-risk group definition excluded HPV 16 infection, indicating that this observation was not influenced by the high number of HPV 16 positives. In both cases, the poor agreement for detection of low-risk HPV genotypes was largely attributable to a higher proportion of low-risk types detected only by the self-collected swab [McNemar’s χ^2 for the ectocervix *versus* self sample = 10.89 ($P = 0.001$) and 0.80 ($P = 0.37$) for low and high risk, respectively; McNemar’s χ^2 for the endocervix *versus* self-sample = 10.89 ($P = 0.001$) and 1.20 ($P = 0.27$) for low and high risk, respectively]. When the ectocervical and endocervical swab results were combined into a single, clinician administered HPV result, the difference in agreement stratified by low-risk HPV types ($\kappa = 0.66$) and high-risk HPV types ($\kappa = 0.78$) was decreased ($t = -1.45$; $P = 0.15$; Table 2). However, the discrepant results among the low-risk HPV genotypes were still more likely to be detected by the self-administered swab only (McNemar’s $\chi^2 = 6.25$, $P = 0.01$ *versus* $\chi^2 = 1.5$; $P = 0.22$ for high-risk genotypes).

We examined possible confounding of agreement by age by estimating age-stratified correlations (Table 3) in 10-year strata (<30, 30–39, 40–49, 50–59, and 60+). There was a suggestion of decreased agreement among the women >60; however, the confidence limits around the κ estimate were broad. To examine the possibility that the decrease in agreement in women >60 years was because of their being postmenopausal, we looked at agreement by menopausal status and found no difference in agreement between pre- and postmenopausal women ($\kappa = 0.72$ *versus* 0.75 ; $t = -0.16$; $P = 0.87$). We restricted the menopausal stratification to women in the 50–59-year age group, where there was adequate representation of both pre- and postmenopausal women, and found no significant differences in agreement (data not shown).

When stratifying by case-control status, the agreement between clinician-administered and self-administered swab results did not change from the overall estimate of agreement ($\kappa = 0.66$ for both case and control strata). We further stratified the case group by histological makeup (*i.e.*, adenocarcinoma *versus* SCC) and again found no statistically significant difference in agreement between the collection methods ($\kappa = 0.64$ for SCC *versus* 0.69 for adenocarcinoma; $t = -0.22$; $P = 0.83$).

Samples were collected from six different clinical centers, and we examined agreement among the three larger sites. We found no difference in agreement when stratifying by these three sites, with κ ranging from 0.72 to 0.80 .

Most samples were collected on the same day; however, a total of 29 subjects collected the self-administered samples on a different day from the clinic visit (median difference of 1 day; range, self-collection 1 year prior to 25 days after clinician collection; 60% collected within the same week). When stratifying between samples that were and were not collected on the

same day, we found evidence for a reduction in agreement when the swabs were collected on different days, but this difference was not significant ($\kappa = 0.76$ for same day *versus* 0.51 for different day collection; $t = 1.29$; $P = 0.20$).

Data on the attitude of the women asked to participate with regard to acceptability of self-collection of samples for HPV testing was not collected in this study. However, there is some indirect evidence to suggest that the option of self-collection would improve the overall participation among the control women. Of the 307 control women who were interviewed and asked to submit a cell sample for HPV testing, 83% agreed to submit a self-collected sample, whereas only 51% agreed to a clinician-collected sample for HPV testing ($t = 11.2$; $P \ll 0.001$). In contrast, there was no difference in participation rates by collection method among the cases. Of the 458 eligible cases (SCC and adenocarcinomas combined), 245 (53.4%) submitted a self-administered sample, and 246 (53.7%) submitted a clinician-administered sample ($t = -0.086$; $P = 0.9316$).

Discussion

HPV DNA detection from self-collected cervicovaginal cells in this study is highly comparable with cervix-directed cell collection performed by a trained clinician. This finding is in agreement with smaller studies that have evaluated various forms of cervicovaginal cell collection for HPV DNA testing (7–11). Additionally, by comparing endocervical and ectocervical clinician-administered samples, we show that endocervical infections (like those giving rise to adenocarcinoma) are also detected using self-collection method, albeit with slightly lower agreement than the ectocervical *versus* self-collected swabs. Here, we used a relatively simple method for collection that does not require special processing for DNA extraction and is amenable to various DNA detection assays. The self-collected samples were equally suitable for DNA testing, as determined by the high percentage of samples positive for β -globin amplification (99.7%; same as either clinician-administered sample). These results are important in addressing the feasibility of self-collection for HPV DNA testing in natural history studies where low response rates are frequently attributable to refusal or reluctance of participants to agree to clinic visits for pelvic examinations and collection of cervical cells.

Although the agreement between the two collection methods was good, each method appeared to miss approximately 10–17% of the positive samples. The discordance of HPV positivity likely reflects fluctuation of DNA detection in samples with viral DNA quantities near the detection limit of the assay. In studies evaluating the association of HPV with cervical cancer, these fluctuations in detectability are unlikely to significantly affect the risk estimates. However, in natural history studies of HPV, involving repeated measures of HPV DNA status over time, these results suggest that multiple swab samples collected from each visit may be required to detect these low-level infections.

There may be some question regarding the generalizability of the results from this population to a more screening-like population that would be the basis for most natural history studies, where self-sampling is likely to be used. Although this is a valid concern, we feel that because the correlation did not differ between the controls (obtained through random-digit dialing) and either case type in our study, the results are equally generalizable to women with and without cervical disease. It is, in fact, reassuring that the control women had equally good agreement between the self- and clinician-administered swab results compared with cases, because the control women are likely to have lower levels of HPV DNA and represent the target population to which self-sampling is most likely to be directed. The women

participating in our study were older on average (*versus* a random population sample) because of the age-matching criteria used in the case-control design, with only 25% <30 years of age. However, stratification by age did not significantly change the correlation between swab samples, suggesting that the age of the population would not affect the ability to gain accurate HPV information from self-collected cervicovaginal cells.

This study has shown strong evidence of the comparability of HPV detection between self-collected and clinician-collected cell samples among women who agreed to participate. Control women, who would be most likely to refuse because of the requirement for a clinic visit, were significantly more likely to agree to submit a self-collected swab *versus* a clinician-administered swab. The difference in response occurred despite an additional monetary incentive and travel reimbursement offered to participants consenting to a clinician-collected swab sample. The difference in response rates indicates an anticipated increase in participation when self-collection for HPV testing is presented as an option. This is in agreement with several other studies that have directly measured the preference of self-collection to clinician-collection (22). Furthermore, because the agreement did not differ by clinic site, with multiple staff members recruiting participants and instructing in self-collection, we might infer practical feasibility of a broad application of the self-sampling method.

Other investigators interested in self-collection of cervicovaginal cells for HPV detection have used a variety of collection instruments including tampons, swabs, and cervicovaginal lavages. We chose to use a Dacron swab as the self-collection instrument in this study for several reasons: (a) we excluded the MY-PAP (a self-administered cervicovaginal lavage) as a collection instrument because it is more difficult to use than either the swab or tampon; (b) we felt that the size and ease of manipulation of the swab would help to ensure compliance; and most importantly, (c) the swab sample does not require the special procedures for DNA extraction that are required for tampon-collected cells. In fact, the self-collected swab sample is processed identically to the clinician-administered swabs and is amenable to a variety of DNA detection methods including L1 consensus PCR and the Food and Drug Administration-approved Digene Hybrid Capture method. A previous study using swabs as the self-collection device and a less sensitive dot-blot method for the detection of HPV DNA showed good correlation with clinician-administered swabs (91% agreement; Ref. 7). We have shown similar success in swab self-collection using a more sensitive, PCR-based DNA detection method, further substantiating the comparability of the self- and clinician-collected samples. In addition to the utility of self-collected swabs for HPV DNA testing, other groups have shown similarly good correlation with clinician-administered swabs in detecting other common sexually transmitted agents, including *Chlamydia trachomatis* and Group B *streptococcus* (22–24).

Previous studies that evaluated two collection methods used randomized designs to assess the effect of sample order in the correlation of HPV detection. We encouraged all women to do the self-collection procedure before the pelvic examination and did not document deviations from this suggestion. The increase in total positive samples seen when comparing the self-administered sample to either clinician-administered swab may arguably represent an effect of sampling order, where more exfoliated cells were collected during the first, self-administered swab sample. However, the fact that the increase in positivity with the self-administered method is largely restricted to low-risk HPV genotypes suggests that the self-collected swab sample collected a larger proportion of vaginal/

vulvar cells that may be more likely to be infected with these low-risk genotypes. A recent small study comparing self-collection by tampon with clinician collection by swab assessed the effect of collection order by randomizing the order of tampon and swab collection. That study saw a trend toward increased positivity in the swabs that were collected first for grouped HPV results, but no such trend of order effect when restricting the analysis to high-risk positive HPV types (10). Their results may also reflect more vaginal/vulvar collection of low-risk types during the first swab collection.

Our original results for the detection of HPV 16, the most prevalent HPV genotype in this and most studies of HPV, were suspect because of a disproportionately high prevalence among the controls and the lack of correlation with sexual behavior. We believe that we isolated the source of the apparent contamination to a single lot of STM that was presumably contaminated with HPV 16 plasmid during the manufacturing process. To resolve true HPV 16 infection from HPV 16 plasmid contamination, we used a unique set of anticontamination primers that amplified only wild-type HPV 16 by spanning the cloning site. The proportion of HPV 16-positive samples from this second amplification was markedly reduced for the suspected STM lot but not the other STM lots, thus providing evidence that one particular lot of STM contained contaminating HPV DNA. We resolved HPV 16-positive infection as those samples positive using the anticontamination primers. There is likely residual misclassification of HPV 16 status because of the knowledge that the anticontamination primers are slightly less sensitive than the L1 consensus primer system for the detection of native HPV 16. Therefore, some of the disagreement between the different sampling methods involving HPV 16 may be an artifact of the decreased sensitivity of the HPV 16 detection method. However, this is unlikely to contribute substantially, because the agreement among HPV 16-positive results did not differ substantially from that of the other HPV genotypes.

We have demonstrated the technical feasibility of HPV DNA testing using cells obtained from self-collected swab material. The absolute agreement by HPV status and HPV type were both good, supporting the utility of self-collection of cervicovaginal cells for natural history studies of HPV. Furthermore, the high percentage of agreement among high-risk HPV genotypes suggests that self-sampling may also provide a reasonable alternative for HPV testing in natural history studies of cervical cancer. This method appears to be more acceptable to women and should decrease the cost associated with HPV testing in investigational studies by eliminating the need for a clinic visit to conduct a full pelvic examination.

Acknowledgments

We are grateful to Barbara Moscicki (University of California-San Francisco, San Francisco, CA) for sharing her experience using self-collection by Dacron swab, and we thank Mark Schiffman (National Cancer Institute, Bethesda, MD) for critical manuscript review. We thank the field study coordinators at Westat, Inc., Jeanne Rosenthal, Shirley Friend, Pat Clark, and Beth Mittl; Kay Helgesen at IMS, Inc.; Shelley Niwa at Westat, Inc., who prepared data for analysis; Tom Alessi at Roche Molecular Systems, who assisted in the HPV PCR assay; Sue Bedger (Hershey Medical Center, Hershey, PA), Michelle Blanchard (Georgetown University, Washington, DC), Lynn Crawford (University of Maryland, Baltimore, MD), Kate Nellemann (George Washington University, Washington, DC), and Bobbi Robbins (Graduate Hospital, Philadelphia, PA), who coordinated efforts at the clinical centers; and Robert Kurman (Johns Hopkins Medical Institutions, Baltimore, MD) and Steven Silverberg (University of Maryland), who along with Richard Zaino (Hershey Medical Center), were responsible for the pathological review of cases.

References

1. Schiffman, M. H., Bauer, H. M., Hoover, R. N., Glass, A. G., Cadell, D. M., Rush, B. B., Scott, D. R., Sherman, M. E., Kurman, R. J., and Wacholder, S.

Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J. Natl. Cancer Inst.* 85: 958-964, 1993.

2. Liaw K-L, Hsing, A. W., Chen C-J, Schiffman, M. H., Zhang, T. Y., Hsieh C-Y, Greer, C. E., You, S-L, Huang, T. W., Wu, T-C., O'Leary, T. J., Seidman, J., Blot, W. J., Meinert, C. L., and Manos, M. M. Human papillomavirus and cervical neoplasia: a case-control study in Taiwan. *Int. J. Cancer*, 62: 565-571, 1995.

3. Koutsky, L. A., Holmes, K. K., Critchlow, C. W., Stevens, C. E., Paavonen, J., Beckmann, A. M., DeRouen, T. A., Galloway, D. A., Vernon, D., and Kiviat, N. B. A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. *N. Engl. J. Med.*, 327: 1272-1278, 1992.

4. Remmick, A. J., Walboomers, J. M. M., Helmerhorst, T. 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.

5. Ho, G. Y., Bierman, R., Beardsley, L., Chang, C. J., and Burk, R. D. Natural history of cervicovaginal papillomavirus infection in young women. *N. Engl. J. Med.*, 338: 423-428, 1998.

6. Liaw K-L, Glass, A. G., Manos, M. M., Greer, C. G., Scott, D. R., Sherman, M. E., Burk, R. D., Kurman, R. J., Wacholder, S., Rush, B. B., Cadell, D. M., Lawler, P., Tabor, D., and Schiffman, M. Detection of human papillomavirus DNA in cytologically normal women and subsequent cervical squamous intraepithelial lesions. *J. Natl. Cancer Inst.*, 91: 954-960, 1999.

7. Moscicki, A-B. Comparison between methods for human papillomavirus DNA testing: a model for self-testing in young women. *J. Infect. Dis.*, 167: 723-725, 1993.

8. Fairley, C. K., Chen, S., Tabrizi, S. N., Quinn, M. A., McNeil, J. J., and Garland, S. M. Tampons. A novel patient-administered method for the assessment of genital human papillomavirus infection. *J. Infect. Dis.*, 165: 1103-1106, 1992.

9. Coutlee, F., Hankins, C., Lapointe, N., and The Canadian Women's HIV Study Group. Comparison between vaginal tampon and cervicovaginal lavage specimen collection for detection of human papillomavirus DNA by the polymerase chain reaction. *J. Med. Virol.*, 51: 42-47, 1997.

10. Harper, D. M., Hildesheim, A., Cobb, J. L., Greenberg, M., Vaught, J., and Lorincz, A. T. Collection devices for human papillomavirus. *J. Fam. Pract.*, 48: 531-535, 1999.

11. Morrison, E. A. B., Goldberg, G. L., Hagan, R. J., Kadish, A. S., and Burk, R. D. Self-administered home cervicovaginal lavage: a novel tool for the clinical-epidemiologic investigation of genital human papillomavirus infections. *Am. J. Obstet. Gynecol.*, 167: 104-107, 1992.

12. Gravitt, P. E., Peyton, C. L., Apple, R. J., and Wheeler, C. M. Genotyping of 27 human papillomavirus types by using L1 consensus PCR products by a single-hybridization, reverse line blot detection method. *J. Clin. Microbiol.*, 36: 3020-3027, 1998.

13. Lacey, J. V., Jr., Brinton, L. A., Abbas, F. M., Barnes, W. A., Gravitt, P. E., Greenberg, M. D., Greene, S. M., Hadjimichael, O. C., McGowan, L., Mortel, R., Schwartz, P. E., Silverberg, S. G., and Hildesheim, A. Oral contraceptives as risk factors for cervical adenocarcinomas and squamous cell carcinomas. *Cancer Epidemiol. Biomark. Prev.*, 8: 1079-1085, 1999.

14. Bauer, H. M., Greer, C. E., and Manos, M. M. Determination of genital human papillomavirus infection using consensus PCR. In: C. S. Herrington and J. O. D. McGee (eds.), *Diagnostic Molecular Pathology: A Practical Approach*, pp. 132-152. Oxford, United Kingdom: Oxford University Press, 1992.

15. Bosch, F. X., Manos, M. M., Munoz, N., Sherman, M., Jansen, A. M., Peto, J., Schiffman, M. H., Moreno, V., Kurman, R., and Shah, K. V. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J. Natl. Cancer Inst.*, 87: 796-802, 1995.

16. van den Brule, A. J., Class, E. C., du Maine, M., Melchers, W. J., Helmerhorst, T., Quint, W. G., Lindeman, J., Meijer, C. J., and Walboomers, J. M. Use of anti-contamination primers in the polymerase chain reaction for the detection of human papillomavirus genotypes in the cervical scrapes and biopsies. *J. Med. Virol.*, 29: 20-27, 1989.

17. Fleiss, J. H. (ed.). The measurement of interrater agreement. In *Statistical Methods for Rates and Proportions* (Second Edition), John Wiley & Sons, New York, NY, pp. 211-236, 1981.

18. Greenland, S. Applications of Stratified Analysis Methods. In: K. J. Rothman and S. Greenland (eds.), *Modern Epidemiology*, Ed. 2, pp. 253-280. Philadelphia: Lippincott-Raven, 1998.

19. Greenland, S. Analysis of Polytomous exposures and outcomes. In: K. J. Rothman and S. Greenland (eds.), *Modern Epidemiology*, Ed. 2, pp. 301-328. Philadelphia: Lippincott-Raven, 1998.

20. Stata Reference Manual, Release 6, Volume 4, pp. 112-119. College Station, TX: Stata Press, 1999.

21. StataCorp. Stata Statistical Software: Release 6.0. College Station, TX: Stata Corporation, 1999.

22. Molnar, P., Biringer, A., McGeer, A., McIsaac, W., and The Mount Sinai GBS Screening Group. Can pregnant women obtain their own specimens for Group B Streptococcus? A comparison of maternal versus physician screening. *Fam. Pract.*, 14: 403-406, 1997.

23. Wiesenfeld, H. C., Heine, R. P., Rideout, A., Macio, I., DiBiasi, F., and Sweet, R. L. The vaginal introitus: a novel site for *Chlamydia trachomatis* testing in women. *Am. J. Obstet. Gynecol.*, 174: 1542-1546, 1996.

24. Domeika, M., Bassiri, M., Butrimiene, I., Venalis, A., Ranceva, J., and Vasjanova, V. Evaluation of vaginal introital sampling as an alternative approach for the detection of genital *Chlamydia trachomatis* infection in women. *Acta Obstet. Gynecol. Scand.*, 78: 131-136, 1999.