

Epidemiology of Acquisition and Clearance of Cervical Human Papillomavirus Infection in Women from a High-Risk Area for Cervical Cancer

Eduardo L. Franco,¹ Luisa L. Villa,⁴ João P. Sobrinho,⁴ José M. Prado,⁴ Marie-Claude Rousseau,¹ Marie Désy,² and Thomas E. Rohan³

¹Departments of Oncology and Epidemiology, McGill University, and ²University of Quebec, Montreal, and ³Department of Public Health Sciences, University of Toronto, Canada; ⁴Ludwig Institute for Cancer Research, São Paulo, Brazil

Acquisition and clearance of cervical human papillomavirus (HPV) infection were analyzed among 1425 low-income women attending a maternal and child health program in São Paulo, Brazil. Specimens collected every 4 months were tested by a polymerase chain reaction protocol (MY09/11). In all, 357 subjects were positive at least once. There were 1.3% new infections per month, with 38% cumulative positivity after 18 months. Of 177 positive subjects at enrollment, only 35% remained infected after 12 months. The monthly clearance rate was higher for nononcogenic types (12.2%; 95% confidence interval [CI], 9.6–15.4) than for oncogenic HPV infections (9.5%; 95% CI, 7.5–11.9). Median retention times were 8.1 months (95% CI, 7.8–8.3) for oncogenic types and 4.8 months (95% CI, 3.9–5.6) for nononcogenic HPV infections. The mean infection durations were 8.2 and 13.5 months for nononcogenic and oncogenic types, respectively. Although a woman's age did not affect mean duration for oncogenic types (13–14 months), nononcogenic-type infections lasted longer (10.2 months) among younger (<35 years old) than in older women (5.6 months).

The epithelial lining of the anogenital tract is the target for infection by a group of mucosotropic viruses, the human papillomaviruses (HPV). Acquisition of cervical HPV infection is the main biologic precursor of a series of events that leads to cervical cancer, as has been extensively documented by epidemiologic and experimental studies during the past 15 years [1, 2].

In cross-sectional surveys, cervical HPV infection is detected by amplified and nonamplified DNA detection techniques in 5%–40% of asymptomatic women of reproductive age [3, 4]. A few cohort studies indicate that HPV infection is mostly a transient or intermittent phenomenon; only a small proportion of those positive for a given HPV type tend to harbor the same type in subsequent specimens [5–8]. In addition, prospective epidemiologic studies show that the risk of subsequent cervical intraepithelial neoplasia seems to be proportional to the num-

ber of specimens testing positive for HPV [9, 10], which suggests that only persistent infections may trigger carcinogenic development. Given that there is now considerable interest in the possibility of using HPV testing as a potential cervical cancer screening tool [11–13], it is imperative that issues related to viral persistence be addressed by epidemiologic studies. Research on the epidemiology of viral persistence and of its determinants will help in the future formulation of algorithms and policies for inclusion of HPV testing in cervical cancer prevention.

In 1993, we began the Ludwig-McGill Cohort Study, a large longitudinal investigation of the natural history of HPV infection and cervical neoplasia in a population of low-income women in São Paulo, one of the highest risk areas worldwide for cervical cancer. It focuses on persistent infection with oncogenic HPV types as the precursor event leading to cervical neoplasia and attempts to understand attributes of the natural history of viral infection that may be helpful for designing primary and secondary strategies for preventing cervical cancer. Here we present the descriptive epidemiologic results concerning the dynamics of acquisition and loss of HPV infection during the first four visits for the first 1425 women enrolled in the cohort.

Subjects and Methods

Study design and population. The Ludwig-McGill Cohort Study is an ongoing longitudinal investigation involving repeated measurements of lifestyle, nutritional, and behavioral risk factors as determinants of the acquisition of HPV infection and precursor lesions of cervical cancer. The design and methods of the investi-

Received 22 March 1999; revised 30 June 1999; electronically published 1 October 1999.

Presented in part: 17th International Papillomavirus Conference, Charleston, South Carolina, January 1999 (abstract 81).

Each study participant signed an informed consent document. All study procedures and the informed consent document were approved by the institutional review boards and ethical committees of all institutions with which the authors are affiliated. Human experimentation guidelines of the US Department of Health and Human Services and of the authors' institutions were followed.

Grant support: NIH (CA-70269); Medical Research Council of Canada (MA-13647).

Reprints or correspondence: Prof. E. Franco, Dept. of Oncology, McGill University, 546 Pine Ave. W., Montreal, QC, Canada H2W 1S6 (eduardof@oncology.lan.mcgill.ca).

The Journal of Infectious Diseases 1999;180:1415–23

© 1999 by the Infectious Diseases Society of America. All rights reserved.
0022-1899/1999/18005-0001\$02.00

gation have been described elsewhere [14]. Study participants attend a comprehensive maternal and child health program (Maternidade Escola Vila Nova Cachoeirinha) that serves low-income families in the city of São Paulo. Women were eligible to participate if they met the following criteria: (1) age between 18 and 60 years; (2) permanent residents of São Paulo (city); (3) not pregnant and not intending to become pregnant in the following 12 months; (4) having an intact uterus and no current referral for hysterectomy; (5) no use of vaginal medication in the previous 2 days; and (6) no treatment for or referral for treatment of cervical disease in the previous 6 months. In addition, women were considered ineligible if they were not interested in complying with all scheduled returns for at least the subsequent 2 years.

The cohort investigation began in November 1993 and accrued subjects until March 1997. Subjects will be followed for up to 5 years: every 4 months in the first year and twice yearly thereafter. To encourage compliance with follow-up returns, women were given meal tickets that increased by \$5 (US) in cash value for each return visit subsequent to enrollment up to a maximum of \$20 at the fourth visit. At each visit, subjects complete an interviewer-administered structured questionnaire and have cervical specimens taken for HPV DNA testing and Pap cytology. The present analysis includes 1425 women whose cervical specimens from the enrollment visit and for up to three follow-up returns have been tested for HPV positivity.

HPV testing. An Accelon biosampler (Medscand, Hollywood, FL) was used to collect a sample of ectocervical and endocervical cells in a tube containing Tris-EDTA buffer, pH 7.4. DNA samples were purified by spin column chromatography. Cervical specimens were tested for the presence of HPV DNA by the MY09/11 polymerase chain reaction protocol [6, 15]. Typing of the amplified products was performed by hybridization with individual oligonucleotide probes specific for 27 genital HPV types. Amplified products that hybridize with the generic probe but with none of the type-specific probes were further tested by restriction fragment length polymorphism (RFLP) analysis [16], which extended the range of identifiable HPV types to >40 genital types. To verify the specificity of the hybridizations, we included more than 30 type-specific positive controls in all membranes. To check the integrity of the host DNA material extracted from the specimens, assays also included an additional set of primers (GH20 and PC04) to amplify a 268-bp region of the β -globin gene [15]. All HPV assays were done blindly on coded specimens with no identification to link different specimens from the same woman. Appropriate precautions were taken to reduce the possibility of specimen contamination.

In some analyses, HPV types were grouped by oncogenicity. We used an expanded classification based on Bauer et al. [17], which grouped HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 68 as oncogenic (high risk) and considered all other HPV types as nononcogenic (low risk). Unknown types were included in the latter group.

Statistical analysis. We used Pearson's χ^2 test to compare frequencies of selected characteristics between women with and without HPV test results during follow-up. Presumed incidence rates of infection with individual HPV types and their associated 95% confidence intervals (CIs) were calculated on the basis of the numbers of cases in which a given type was detected among women

who were free of that type at the enrollment visit. Follow-up time was considered to be until the first visit in which the stated type was detected or until the last visit if the subject remained negative for that type throughout the observation period. We used the Kaplan-Meier method to calculate both the cumulative probability of HPV detection among women who were HPV negative at enrollment and the proportion of women remaining HPV positive at different times after enrollment. Missing values during follow-up because of lack of a specimen or a β -globin-negative result were treated as censored observations unless there was a subsequent informative specimen at a later visit, in which case the result for the latter was used. The statistical assessment of linear dose-response relations between ordinal variables and outcome was based on the χ^2 test for trend.

We assessed the tendency for infection with certain HPV type combinations to occur more frequently than others by computing expected frequencies of joint cumulative positivity for all pairwise combinations under the assumption of no association between types. Expected values were compared against observed frequencies to screen for types occurring in association. We used Fisher's exact test to gauge statistically significant departures from the expected values among all type pairs. This was done solely to flag possible patterns in type associations and not as a means of formally assessing the significance of the individual associations. CIs for observed to expected (O/E) ratios were based on the Poisson distribution.

Results

Characteristics of the cohort. The 1425 participants in this analysis represented mostly the subcohort of women admitted early in the study (the cohort enrolled 2528 women with a 70% response rate). In total, we tested 4873 cervical specimens from enrollment and follow-up visits (mean, 3.4 specimens/subject). The follow-up experience accumulated for this subcohort was 14,341 woman-months (mean, 10.1 months/participant; range, 3.6–36.7). Specimens from 1067 women (75% of the original subcohort) were available at the fourth visit. The median age was 33 years (mean, 33.3; first and third quartiles, 26 and 39 years, respectively). Only 12% of the subjects were ≥ 45 years old. Most participants (66%) were white, and 83% had at most an elementary school education. Slightly more than 50% of the women reported ever smoking cigarettes; 35% were current smokers. Only 1% of the women never were pregnant; 85% reported at least two pregnancies. Only 36% of the women reported having had >2 lifetime sex partners.

There were 172 subjects for whom HPV results were not available at the second visit (12% of the original subcohort). These women were mostly comparable to the rest of the subcohort with respect to race ($P = .536$), age ($P = .121$), education ($P = .204$), age at first intercourse ($P = .180$), parity ($P = .740$), and duration of oral contraceptive pill use ($P = .154$). On the other hand, those who were not followed up tended to have slightly fewer lifetime sex partners (mean, 2.9) than those who

had HPV results during follow-up (mean, 3.6); the difference was of marginal statistical significance ($P = .075$).

Cumulative HPV positivity. Table 1 shows the prevalence of HPV infection at enrollment and the cumulative HPV positivity for each HPV type and for unknown HPVs. Of the genital HPV types that can be specifically ascertained by typing, only HPV-64 was not detected. The overall prevalence of infection with any HPV type at enrollment was 13.8% (196 positive cases). The point prevalence of HPV infection at each subsequent follow-up visit varied little: 14.9%, 14.7%, and 12.3% for visits 2–4, respectively. The prevalence of multiple-type infections (2–4 types) was 2.2% at enrollment and ranged between 1.2% and 1.7% in subsequent visits (data not shown). A total of 357 subjects were positive at least once for HPV, either at the enrollment visit or subsequently, which yielded a crude (nonactuarial) cumulative rate of 25.1%. HPV-16 was the

most common type cumulatively ($n = 60$, 4.2%), followed by HPV-53 ($n = 43$, 3.0%), unknown types ($n = 38$, 2.7%), HPV-58 ($n = 32$, 2.2%), and HPV-31 ($n = 28$, 2.0%; table 1).

Table 1 also classifies HPV types on the basis of the tendency for being persistently detected during follow-up visits. The right-hand column in table 1 shows the ratio between the frequency of positivity in two or more visits and that in a single visit as indicative of the tendency for a given type to persist. The ratio for most HPV types was <1.0 . The highest ratios were those for HPV-33, -40, and -61, although they were based on small individual frequencies. The ratio for HPV-16 was 0.9, indicating that an about equal number of subjects had positivity in two or more visits, compared with positivity only once.

Presumed incidence of HPV infection. Acquisition of infection by any HPV type was relatively common: 1.3%/month (95% CI, 1.1–1.6). Table 2 shows the presumed incidence rate

Table 1. Prevalence of human papillomavirus (HPV) genotypes at enrollment and cumulative positivity in up to three follow-up visits.

HPV type	Prevalence at entry (%)	Cumulative positivity (%)	Positive only once (%)	Positive ≥ 2 visits (%)	Ratio $\geq 2/\text{once}$
6/11	14 (1.0)	25 (1.8)	19 (1.3)	6 (0.4)	0.3
16	38 (2.7)	60 (4.2)	31 (2.2)	29 (2.0)	0.9
18	11 (0.8)	19 (1.3)	11 (0.8)	8 (0.6)	0.7
26	1 (0.1)	3 (0.2)	3 (0.2)	0 (—)	0.0
31	15 (1.1)	28 (2.0)	14 (1.0)	14 (1.0)	1.0
32	1 (0.1)	2 (0.1)	2 (0.1)	0 (—)	0.0
33	6 (0.4)	12 (0.8)	5 (0.4)	7 (0.5)	1.4
34	2 (0.1)	2 (0.1)	2 (0.1)	0 (—)	0.0
35	3 (0.2)	10 (0.7)	7 (0.5)	3 (0.2)	0.4
39	1 (0.1)	5 (0.4)	3 (0.2)	2 (0.1)	0.7
40	4 (0.3)	9 (0.6)	4 (0.3)	5 (0.4)	1.3
42	0 (—)	1 (0.1)	1 (0.1)	0 (—)	0.0
44	0 (—)	1 (0.1)	1 (0.1)	0 (—)	0.0
45	7 (0.5)	14 (1.0)	12 (0.8)	2 (0.1)	0.2
51	10 (0.7)	25 (1.8)	17 (1.2)	8 (0.6)	0.5
52	9 (0.6)	25 (1.8)	18 (1.3)	7 (0.5)	0.4
53	20 (1.4)	43 (3.0)	26 (1.8)	17 (1.2)	0.7
54	1 (0.1)	6 (0.4)	4 (0.3)	2 (0.1)	0.5
55	2 (0.1)	9 (0.6)	8 (0.6)	1 (0.1)	0.1
56	8 (0.6)	14 (1.0)	10 (0.7)	4 (0.3)	0.4
57	0 (—)	1 (0.1)	1 (0.1)	0 (—)	0.0
58	17 (1.2)	32 (2.2)	18 (1.3)	14 (1.0)	0.8
59	1 (0.1)	6 (0.4)	6 (0.4)	0 (—)	0.0
61	7 (0.5)	15 (1.1)	6 (0.4)	9 (0.6)	1.5
62	1 (0.1)	6 (0.4)	4 (0.3)	2 (0.1)	0.5
66	0 (—)	5 (0.4)	3 (0.2)	2 (0.1)	0.7
67	0 (—)	2 (0.1)	2 (0.1)	0 (—)	0.0
68	5 (0.4)	10 (0.7)	9 (0.6)	1 (0.1)	0.1
69	0 (—)	1 (0.1)	1 (0.1)	0 (—)	0.0
70	12 (0.8)	18 (1.3)	10 (0.7)	8 (0.6)	0.8
72	1 (0.1)	3 (0.2)	0 (—)	3 (0.2)	Infinity
73	3 (0.2)	8 (0.6)	6 (0.4)	2 (0.1)	0.3
MM8 (Pap155)	7 (0.5)	20 (1.4)	14 (1.0)	6 (0.4)	0.4
MM7 (Pap291)	2 (0.1)	7 (0.5)	6 (0.4)	1 (0.1)	0.2
MM4 (W13B)	1 (0.1)	4 (0.3)	3 (0.2)	1 (0.1)	0.3
IS39	2 (0.1)	4 (0.3)	4 (0.3)	0 (—)	0.0
C6108	0 (—)	1 (0.1)	1 (0.1)	0 (—)	0.0
C8304	3 (0.2)	10 (0.7)	7 (0.5)	3 (0.2)	0.4
C8061	1 (0.1)	7 (0.5)	6 (0.4)	1 (0.1)	0.2
Unknown	11 (0.8)	38 (2.7)	34 (2.4)	4 (0.3)	0.1
Any types	196 (13.8)	357 (25.1)	202 (14.2)	155 (10.9)	0.8

NOTE. Data are from 1425 subjects with ≤ 4 visits completed and fully typed DNA specimens.

Table 2. Presumed incidence of infection with the 10 most frequently detected individual human papillomavirus (HPV) types and for groups according to oncogenicity.

HPV type	No. of incident cases	No. of women at risk ^a	Woman-months of follow-up	Rate/1000 woman-months (95% confidence interval)
6/11	10	1348	14,130	0.7 (0.3–1.3)
16	19	1324	13,862	1.4 (0.8–2.1)
18	4	1351	14,225	0.3 (0.1–0.7)
31	13	1347	14,153	0.9 (0.5–1.6)
51	15	1352	14,170	1.1 (0.6–1.8)
52	16	1353	14,195	1.1 (0.6–1.8)
53	22	1342	14,066	1.6 (1.0–2.4)
58	13	1345	14,103	0.9 (0.5–1.6)
70	6	1350	14,187	0.4 (0.2–0.9)
MM8 (Pap155)	12	1355	14,215	0.8 (0.4–1.5)
Unknown types	27	1351	14,081	1.9 (1.3–2.8)
Nononcogenic risk types ^b	116	1267	12,722	9.1 (7.5–10.9)
Oncogenic risk types ^c	86	1248	12,718	6.8 (5.4–8.4)
All types	154	1171	11,467	13.4 (11.4–15.7)

^a Women who were negative for HPV of the stated type or group at enrollment.

^b HPV-6/11, -26, -32, -34, -40, -42, -44, -53, -54, -55, -57, -59, -62, -64, -66, -67, -69, -70, -72, -73, MM4, MM7, MM8, CP6108, CP8061, CP8304, IS39, and unknown types.

^c HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, and -68.

of infection for the most common HPV types found in the study. The number of women at risk for each HPV type represents those who were negative for that type at enrollment. For instance, a woman who was positive for HPV-16 at entry would not be counted in the analysis of HPV-16 incidence but would contribute woman-months of follow-up for all other HPVs.

By grouping according to their oncogenicity of all 40 HPV types that have been identified thus far in the cohort (plus unknown HPVs), the combined prevalence at entry for nononcogenic types was somewhat lower than for oncogenic types (7.0% vs. 8.4%, data not shown). However, the combined incidence for the latter (0.68%/month) was significantly lower than for nononcogenic types (0.91%/month), which indicates the tendency for oncogenic types to be more persistent, leading to longer duration of infection and consequently to a higher point prevalence. HPV-53 was the most frequent incident type (0.16%/month), followed by HPV-16 (0.14%/month). Unknown HPVs of ≥ 1 types accounted for the highest number of new infections (0.19%/month).

Age-specific prevalence and incidence. Table 3 separately presents the initial prevalence and presumed incidence rates of infection during follow-up for the most common HPV types for women aged <35 and for those aged ≥ 35 years. The prevalence and the incidence rates were generally higher for younger women for most of the common types, except for HPV-52 (prevalence) and HPV-70 (incidence). A positive relation with age was seen for both prevalence and incidence of unknown types. By considering grouped types, the negative relation with age was maintained for prevalent and incident oncogenic HPVs, whereas the incidence rates of nononcogenic types were comparable for the two age groups despite a higher prevalence for younger women.

Cumulative HPV positivity. Figure 1 shows the cumulative

probability of HPV detection over time for infection of any type, for nononcogenic and oncogenic types, and for HPV-16 in isolation. Among women who were negative at enrollment, the cumulative rates of infection with any types were 7.3% (95% CI, 5.7–8.9), 13.6% (95% CI, 11.5–15.8), and 23.6% (95% CI, 17.2–30.0) at 6, 12, and 18 months, respectively. The cumulative probability of acquiring infection with a nononcogenic HPV type at 18 months (22.4%; 95% CI, 14.7–30.0) was twice that for an oncogenic type (11.2%; 95% CI, 6.8–15.7). The cumulative proportion testing positive for HPV-16 among women

Table 3. Age-specific initial prevalence and presumed incidence rate of infection with the 10 most frequent individual human papillomavirus (HPV) types and for groups according to oncogenicity.

HPV type	Prevalent cases at enrollment (%) (years)		Incident cases (rate/1000 woman-months) (years)	
	<35	≥ 35	<35	≥ 35
6/11	11 (1.4)	3 (0.5)	8 (1.0)	2 (0.3)
16	27 (3.4)	11 (1.9)	13 (1.6)	6 (1.0)
18	8 (1.0)	3 (0.5)	4 (0.5)	0 (—)
31	13 (1.7)	2 (0.4)	11 (1.3)	2 (0.3)
51	8 (1.0)	2 (0.4)	14 (1.7)	1 (0.2)
52	5 (0.6)	4 (0.7)	12 (1.4)	4 (0.7)
53	16 (2.0)	3 (0.5)	15 (1.8)	7 (1.2)
58	15 (1.9)	2 (0.4)	10 (1.2)	3 (0.5)
70	8 (1.0)	4 (0.7)	2 (0.2)	4 (0.7)
MM8 (Pap155)	5 (0.6)	2 (0.4)	8 (1.0)	4 (0.7)
Unknown types	5 (0.6)	6 (1.1)	15 (1.8)	12 (2.1)
Nononcogenic risk types ^a	66 (8.4)	28 (4.9)	67 (9.0)	49 (9.2)
Oncogenic risk types ^b	83 (10.6)	31 (5.4)	64 (8.8)	22 (4.0)
All types	134 (17.1)	56 (9.8)	97 (14.9)	57 (11.5)

^a HPV-6/11, -26, -32, -34, -40, -42, -44, -53, -54, -55, -57, -59, -62, -64, -66, -67, -69, -70, -72, -73, MM4, MM7, MM8, CP6108, CP8061, CP8304, IS39, and unknown types.

^b HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, and -68.

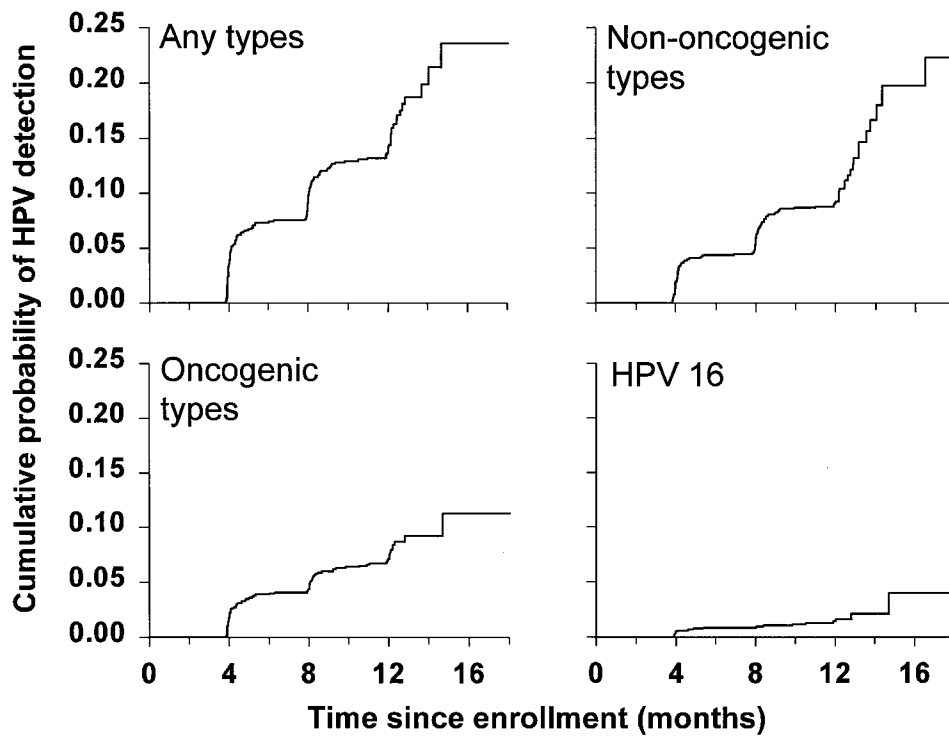


Figure 1. Kaplan-Meier estimates of cumulative human papillomavirus (HPV) positivity for infection with any type, nononcogenic and oncogenic types, and separately for HPV-16. Only women who were negative at enrollment for stated grouping or type were included in analyses for that grouping or type.

who were HPV-16 negative at enrollment was 1.5% at 12 months (95% CI, 0.7–2.2) and 4.0% (95% CI, 0.1–7.8) at 18 months.

Frequency of testing and presumed incidence. We tested the hypothesis that the frequency of return visits was influencing the incidence of HPV infection in the cohort. Although return visits had been prescheduled at 4, 8, and 12 months, a substantial proportion of women delayed their appointments. Once a woman returned, our nurses scheduled the next appointment for 4 months later, which in the end produced sufficient variation in time between visits to allow an analysis of HPV positivity as a function of the number of visits but with adjustment for the person-time experience of follow-up. Table 4 shows that

the frequency of return visits did not influence either the presumed overall HPV incidence rate or the cumulative rate at 12 months (nonsignificant linear trend in both cases). This finding suggests that the detection rates for presumably newly acquired infections are not a function of the intensity of viral surveillance during follow-up insofar as the conditions of our study design are concerned.

Loss of initial positivity. HPV infection was mostly a transient phenomenon in the cohort. Of the 177 subjects who were positive for any type at enrollment and who had one or more follow-up visits with evaluable specimens, 61% (95% CI, 54–69) retained their original positivity at 6 months, but only 35% (95% CI, 27–42) did so after 12 months. Figure 2 shows the

Table 4. Effect of frequency of return visits on presumed incidence rate and cumulative incidence of human papillomavirus (HPV) infection (all types) during follow-up among women who were HPV negative at enrollment.

No. of return visits	No. of subjects	No. of new HPV-positive cases	Woman-months of follow-up	Rate/1000 woman-months (95% CI)	Cumulative incidence (%) at 12 months (95% CI)
1	113	6	538	11.2 (4.1–24.3)	18.0 (3.7–32.4)
2	140	19	1420	13.4 (8.1–20.9)	15.0 (8.1–21.9)
3	805	129	9509	13.6 (11.3–16.1)	13.6 (11.2–15.9)
				<i>P</i> (trend) ^a = .693	<i>P</i> (trend) ^a = .922
All	1058	154	11,467	13.4 (11.4–15.7)	13.6 (11.5–15.8)

NOTE. CI, confidence interval.

^a *P* value from χ^2 statistic for linear trend.

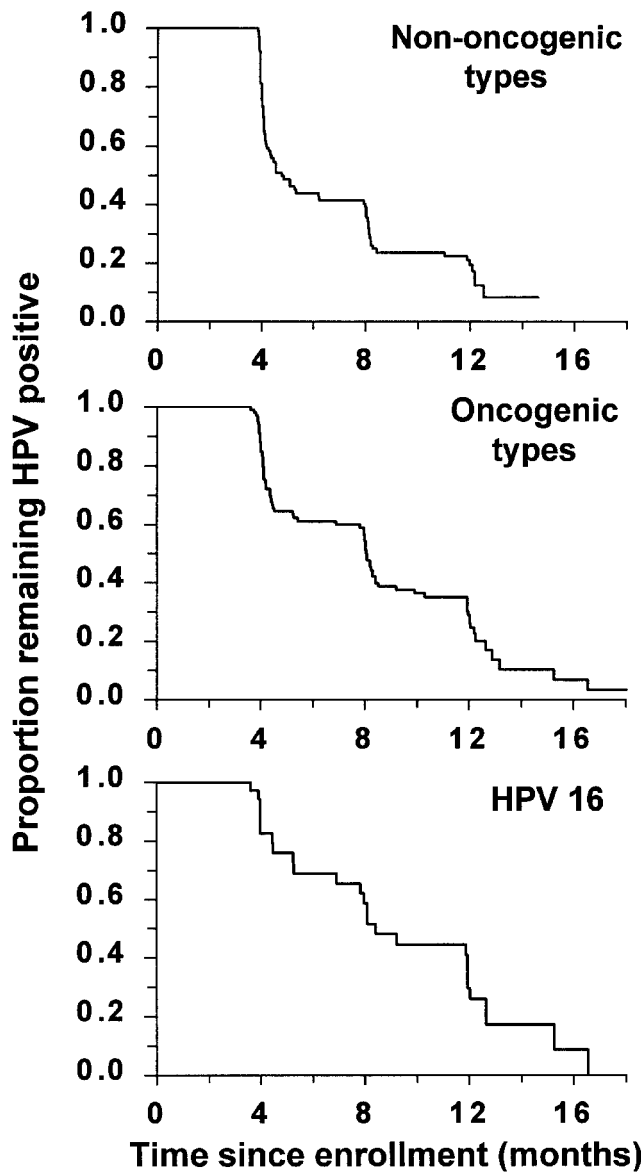


Figure 2. Kaplan-Meier estimates of proportion of women remaining positive for nononcogenic and oncogenic human papillomavirus (HPV) types and for HPV-16. Only women positive at enrollment for stated grouping or type were included in analyses for that grouping or type.

proportions of women remaining positive for specified type combinations and for HPV-16 as a function of time since enrollment. Loss of positivity by group was considered to have occurred at the first follow-up visit at which the subject was no longer positive for the type grouping detected at enrollment, even if the cervical specimen was positive for a viral type of a different grouping. The rate of loss of positivity was higher for nononcogenic types (12.2%/month; 95% CI, 9.6–15.4) than for oncogenic HPV infections (9.5%/month; 95% CI, 7.5–11.9).

This resulted in median positivity retention times since enrollment (i.e., the time interval it took for 50% of the cases to have lost the original positivity) that were substantially greater for oncogenic types (8.1 months; 95% CI, 7.8–8.3) than for nononcogenic types (4.8 months; 95% CI, 3.9–5.6). The mean duration of infections (measured actuarially) detected at enrollment was 7.0 months (95% CI, 6.2–7.8) for nononcogenic and 8.9 months (95% CI, 7.6–10.2), for oncogenic types. Persistence for HPV-16 infections was comparable to that of other oncogenic types: 8.9%/month rate of loss (95% CI, 5.8–13.1), with a median positivity retention of 8.4 months (95% CI, 6.8–10.0).

Joint infections. Table 5 shows all pairwise frequency combinations of HPV types as detected cumulatively since enrollment. Each observed frequency of joint cumulative positivity is contrasted with the expected value under the assumption that no association existed between the 2 types being compared. Only the 13 most common types (those with $\geq 1\%$ cumulative positivity) were included in the analysis. Although unknown HPVs were a frequent finding, they were not included because they are detected by exclusion, which invalidates the analysis of joint frequencies with other types. In all, with 78 possible pairwise combinations, one would expect that for a few type pairs, the observed frequencies would be deemed significantly different from the expected ones by chance alone. Assuming a completely random distribution of the 13 types shown in table 5, one would expect to see about 4 associations exceeding the 5% significance level and only 1 exceeding the 1% level. However, there were 13 pairs at the 5% level, 4 of which had differences above the 1% level. In all instances, the observed value was greater than the expected. Five of these joint excess occurrences involved HPV-53, and, in another 4, HPV-16 was involved; these are the 2 most common types. These 2 types did not seem to be strongly associated (O/E ratio, 2.2; 95% CI, 0.6–5.7). Of interest, the highest frequency of joint positivity involving HPV-16 was with MM8 (O/E ratio, 6.3; 95% CI, 2.0–14.6). HPV-6/11 was only remarkably frequent in association with HPV-53 (O/E ratio, 5.0; 95% CI, 1.4–12.8).

Discussion

Before we address the implications of this study's findings, it is important that we consider its limitations. First, cohort participants were not selected from a probabilistic sampling base from the general population of women at risk for cervical cancer. For logistic and practical reasons, we chose a clinic-based design that selected women who were already clients of a maternal and child health delivery system maintained by the municipality of São Paulo for low-income residents in the northern part of the city. Cohort members were, therefore, not representative of the area's general female population in the eligible age range for the study (18–60 years), who presumably may not have had comparable access to a public health maintenance program. One indication that this may have been so

Table 5. Observed and expected frequencies of joint cumulative positivity for the most common human papillomavirus (HPV) types among all 1425 women enrolled in the cohort.

HPV type	Observed and expected pairwise frequencies of joint cumulative infection											
	16	18	31	45	51	52	53	56	58	61	70	MM8
6/11	1 (1.1)	0 (0.3)	1 (0.5)	1 (0.2)	2 (0.4)	2 (0.4)	4** (0.8)	0 (0.2)	2 (0.6)	2 (0.3)	1 (0.3)	0 (0.4)
16		3* (0.8)	4* (1.2)	2 (0.6)	1 (1.1)	1 (1.1)	4 (1.8)	3* (0.6)	3 (1.3)	1 (0.6)	1 (0.8)	5** (0.8)
18			0 (0.4)	0 (0.2)	0 (0.3)	1 (0.3)	1 (0.6)	1 (0.2)	1 (0.4)	1 (0.2)	0 (0.2)	1 (0.3)
31				1 (0.3)	3* (0.5)	0 (0.5)	3* (0.8)	1 (0.3)	2 (0.6)	0 (0.3)	0 (0.4)	0 (0.4)
45					2* (0.2)	1 (0.2)	1 (0.4)	0 (0.1)	1 (0.3)	1 (0.1)	0 (0.2)	0 (0.2)
51						2 (0.4)	3* (0.8)	0 (0.2)	1 (0.6)	1 (0.3)	1 (0.3)	1 (0.4)
52							4** (0.8)	0 (0.2)	3* (0.6)	0 (0.3)	1 (0.3)	0 (0.4)
53								1 (0.4)	4* (1.0)	1 (0.5)	0 (0.5)	2 (0.6)
56									3** (0.3)	0 (0.1)	0 (0.2)	1 (0.2)
58										1 (0.3)	1 (0.4)	1 (0.4)
61											0 (0.2)	0 (0.2)
70												1 (0.3)

NOTE. Expected frequencies in parentheses. No. of asterisks denotes significance level exceeded (*, $P < .05$; **, $P < .01$).

is that only 5% of our subjects reported never having had a Pap smear [14]. On the other hand, our clinic-based sample was more appropriate for a study of the natural history of HPV infection and cervical cancer precursors because of the likely higher compliance with the requirement for multiple follow-up returns. The study was conducted in a country with a high cervical cancer incidence by international standards [18].

In a study of the natural history of cervical HPV infection, one would expect that estimates of incidence of infections by type and accrued positivity over time would be a function of the intensity of surveillance built into the study design. Subjects were told to return every 4 months during the first year and then every 6 months thereafter for cervical specimen collection. We expected that this relatively intensive regimen would yield artificially high rates of detection and cumulative positivity. Of interest, we found that the number of return visits during follow-up did not influence either the rate of detection of new infections or the cumulative positivity. This suggests that the dynamics of acquisition of infection and its duration are not prone to a surveillance bias related to determinants of compliance to scheduled visits, at least within the constraints of our cohort study design.

Another important caveat in a study of the natural history of HPV infection of women past their onset of sexual activity is that one cannot measure the true incidence of new infections among those negative at enrollment: it is impossible to distinguish a new infection from a recurrence or reactivation of a previously latent infection within the limitations of the molecular sensitivity of the HPV detection method. For this reason, we can only claim to have measured the “presumed” incidence rate of a new type for those negative at enrollment. Had we made our definition more stringent, by including only women with 2 negative HPV results at the enrollment and first follow-up visits in the computation of presumed incidence rates, the results might have been different.

Our cohort included relatively older women compared with some previous studies of the dynamics of HPV infection [7, 8,

19, 20]. Proportionally fewer women in our study had had multiple sex partners compared with other cohorts or cross-sectional surveys [15, 17, 21, 22]. Nevertheless, HPV infection was relatively frequent in the cohort: 38% of the women had ≥ 1 HPV-positive specimen after 18 months. HPV-16 was the most common type (19% of all positive cases at enrollment), not unlike most previous studies (reviewed by [2]). HPV-53 was the second most commonly detected type: 3% cumulative positivity (10% of all positive cases at enrollment), exceeding HPV-6/11, which is typically the second most common in North American studies [2] and which was the second most common in a previous study we conducted in northeast Brazil [23]. Infection with HPV-58 was the third most frequent occurrence in the cohort (2.2% cumulative positivity). This type was highly prevalent in a comparable low-income, mixed ethnicity female population in Washington, DC [21]. Unknown HPVs were a relatively common finding, but their occurrence is an underestimate because multiple infections with known and unknown types are impossible to distinguish in practice because unknown types are detected by exclusion (i.e., when the generic probe is positive in the absence of hybridization with any of the specific probes for individual types or no discernible RFLP pattern). If both the generic and a specific probe were positive (the majority of the HPV-positive specimens), then the result was counted toward that specific type.

The relation between incidence and prevalence is mediated by the average duration of HPV infection, an indicator of persistence, which is a key prognostic variable for risk of high-grade cervical lesions [10, 19]. Using actuarial methods to analyze persistence of infections present at enrollment, we found that positivity for oncogenic types was longer lasting than for nononcogenic types. It took almost twice as long for 50% of the oncogenic-type infections to clear (8.1 months) than for the nononcogenic types (4.8 months).

Calculating mean infection retention times actuarially does not provide a true measure of the average duration of infections because it is impossible to know how long a woman has been

infected by the time she is found positive at enrollment. The resulting mean retention time is likely to underestimate the true mean duration of infection, with the extent of the bias being a function of how persistent a given type may be. One way to assess this discrepancy is by use of an alternative and unbiased measure of mean duration of infection that is based on the relation between prevalence and incidence estimates.

Assuming constant incidence and using the simple formula (derived from [24]): $\text{Duration} = \text{Prevalence} / [\text{Incidence} \times (1 - \text{Prevalence})]$, we obtain 8.2 and 13.5 months as the mean durations for infections with nononcogenic and oncogenic types, respectively. Although both of these estimates are greater than the equivalent actuarial ones, the magnitude of the difference is more appreciable for oncogenic types (13.5 vs. 8.9 months) than for nononcogenic types (8.2 vs. 7.0 months), which underscores the extent of the bias as a function of persistence.

Age seemed to mediate the relation between prevalence and incidence when considering grouped types (i.e., oncogenic and nononcogenic HPVs). An inverse relation with age was seen with both prevalent and incident oncogenic HPVs, whereas the incidence of nononcogenic types did not vary by age, despite the higher prevalence among younger women. This finding suggests that age may not be a factor in determining persistence of oncogenic types, but younger women (<35 years) may be more likely to have longer-lasting infections with nononcogenic types, which would result in a higher prevalence than for older women, despite a comparable incidence as in older women. In fact, by using the above formula separately for the 2 age groups (<35 and ≥ 35 years), we observe that oncogenic-type infections have comparable mean durations by age: 13.4 months for women aged <35 years and 14.2 months for those aged ≥ 35 years, whereas younger women have nononcogenic-type infections that last about twice as long (10.2 months) as those in older women (5.6 months). This interaction effect between HPV type and age on persistence needs to be corroborated by further studies before any firm conclusions can be drawn.

We investigated whether the risk of acquiring multiple infection by specific HPV types was random in the cohort. We anticipated that joint infections would be more common than is expected by chance because of the shared risk determinants among types (e.g., sexual activity, age). Our goal was to identify distribution clusters that could reveal a tendency for certain types to occur more frequently as joint infections, thus providing clues as to the existence of separate transmission patterns. Although our search may have been affected by low statistical power because of the relative rarity of individual-type infections, we found 13 pairs of types with significantly higher than expected joint occurrence. HPV-16 and HPV-53 were frequently found in association with other types but were not remarkably common with each other. Some types were more frequently found in isolation (e.g., HPV-61 and HPV-70) with a tendency for lower than expected joint occurrences. We caution against drawing any inferences concerning possible commonality of

transmission factors, however, because these distribution inequalities could be due to chance or to low statistical power. On the other hand, it is clear that HPV types tend to occur as joint infections far more frequently than one would expect. As our cohort increases in size and follow-up is extended, such patterns could reveal important clues about determinants of transmission for individual HPV types.

Understanding the epidemiology of cervical HPV infection is an important first step toward the development of strategies for preventing genital HPV disease and, ultimately, cervical cancer. Longitudinal studies, such as the present one, are ongoing in different geographic areas and may provide critical information to assist in the implementation of improved screening programs and in the delivery of future immunization programs aiming at preventing HPV infection and cervical neoplasia.

Acknowledgments

We are indebted to Maria L. Baggio and Lenice Galan for patient management and specimen collection, to Antonio L. Ruiz and Marcella P. Ribeiro for laboratory assistance, and to Silvaneide Ferreira for computer data entry.

References

1. IARC Working Group. Human papillomaviruses. IARC Monogr Eval Carcinog Risks Hum **1995**;64:279–82.
2. Schiffman MH, Burk RD. Human papillomaviruses. In: Evans A, Kaslow R, eds. *Viral infections of humans, epidemiology and control. Part III. Viral infections and malignant diseases*. New York: Plenum, **1997**: 983–1023.
3. Franco EL, Villa LL, Richardson H, Rohan T, Ferenczy A. Epidemiology of cervical human papillomavirus infection. In: Franco EL, Monsonogo J, eds. *New developments in cervical cancer screening and prevention*. Oxford, UK: Blackwell Science, **1997**:14–22.
4. Schiffman MH, Brinton LA. The epidemiology of cervical carcinogenesis. *Cancer* **1995**;76:1888–901.
5. Moscicki AB, Palefsky J, Smith G, Siboshki S, Schoolnik G. Variability of human papillomavirus DNA testing in a longitudinal cohort of young women. *Obstet Gynecol* **1993**;82:578–85.
6. Hildesheim A, Schiffman MH, Gravitt PE, et al. Persistence of type-specific human papillomavirus infection among cytologically normal women. *J Infect Dis* **1994**;169:235–40.
7. Hinchliffe SA, VanVelzen D, Korporeal H, Kok PL, Boon ME. Transience of cervical HPV infection in sexually active young women with normal cervicovaginal cytology. *Br J Cancer* **1995**;72:943–5.
8. Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD. Natural history of cervicovaginal papillomavirus infection in young women. *N Engl J Med* **1998**;338:423–8.
9. Koutsky LA, Holmes KK, Critchlow CW, et al. A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. *N Engl J Med* **1992**;327:1272–8.
10. Ho GYF, Burk RD, Klein S, et al. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. *J Natl Cancer Inst* **1995**;87:1365–71.
11. Johnson K. Periodic health examination, 1995 update 1: screening for human papillomavirus infection in asymptomatic women. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J* **1995**;152:483–93.

12. Richart RM. Screening—the next century. *Cancer* **1995**;76:1919–27.
13. Jenkins D, Sherlawjohnson C, Gallivan S. Can papilloma virus testing be used to improve cervical cancer screening. *Int J Cancer* **1996**;65:768–73.
14. Franco EL, Villa L, Rohan T, et al. Design and methods of the Ludwig-McGill longitudinal study of the natural history of human papillomavirus infection and cervical neoplasia in Brazil. *Pan Am J Publ Health* (in press).
15. Bauer HM, Ting Y, Greer CE, et al. Genital human papillomavirus infection in female university students as determined by a PCR-based method. *JAMA* **1991**;265:472–7.
16. Bernard HU, Chan SY, Manos MM, et al. Identification and assessment of known and novel human papillomaviruses by polymerase chain reaction amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms. *J Infect Dis* **1994**;170:1077–85; erratum **1996**;173:516.
17. Bauer HM, Hildesheim A, Schiffman MH, et al. Determinants of genital human papillomavirus infection in low-risk women in Portland, Oregon. *Sex Transm Dis* **1993**;20:274–8.
18. Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J. Cancer incidence in five continents. Vol VII. Lyon, France: IARC Scientific Publications (no. 143), **1997**:1012–3.
19. Moscicki AB, Shiboski S, Broering J, et al. The natural history of human papillomavirus infection as measured by repeated DNA testing in adolescent and young women. *J Pediatr* **1998**;132:277–84.
20. Evander M, Edlund K, Gustafsson A, et al. Human papillomavirus infection is transient in young women—a population-based cohort study. *J Infect Dis* **1995**;171:1026–30.
21. Hildesheim A, Gravitt P, Schiffman MH, et al. Determinants of genital human papillomavirus infection in low-income women in Washington, DC. *Sex Transm Dis* **1993**;20:279–85.
22. Karlsson R, Jonsson M, Edlund K, et al. Lifetime number of partners as the only independent risk factor for human papillomavirus infection—a population-based study. *Sex Transm Dis* **1995**;22:119–27.
23. Franco EL, Villa LL, Ruiz A, Costa MC. Transmission of cervical human papillomavirus infection by sexual activity: differences between low- and high-risk oncogenic types. *J Infect Dis* **1995**;172:756–63.
24. Freeman J, Hutchison GB. Prevalence, incidence and duration. *Am J Epidemiol* **1980**;112:707–23.