

Prospective Study of HPV16 Viral Load and Risk of *In Situ* and Invasive Squamous Cervical Cancer

Karin Sundström¹, Alexander Ploner¹, Lisen Arnheim Dahlström¹, Juni Palmgren^{1,2,7}, Joakim Dillner^{1,3,4}, Hans-Olov Adami^{1,6}, Nathalie Ylitalo^{1,5}, and Pär Sparén¹

Abstract

Background: A strong association has been shown between high viral DNA load (VL) of human papillomavirus (HPV) type 16 and risk for cervical cancer *in situ* (CIS). However, little data is available for the significance of VL in invasive squamous cell carcinoma (SCC).

Methods: In 2 nested case-control studies among women participating in cervical screening, with a cytologically normal first smear, we collected 5,665 smears from 621 women with CIS, 457 with SCC, and individually matched controls. All smears were tested for HPV, and VLs of HPV16 positive smears were quantified using real time-PCR. The median follow-up until diagnosis of CIS or SCC was 6.1 to 7.7 years.

Results: Low VL's were common among both CIS and SCC case women, until 1 to 2 years before diagnosis when a surge in VL occurred. The relative risk (RR) associated with low viral load of HPV16 was around 10 for CIS, and 10 to 20 for SCC throughout 10 years before diagnosis, compared with HPV16-negative women. For women with medium to high VL, the risk for CIS was greatly increased from 5 years before diagnosis [RR, 19; 95% confidence interval (CI), 7–48]. In SCC, a high VL conferred an increased risk, but only from 3 years before diagnosis [RR, 60; 95% CI, 6–580].

Conclusions: We show differing risk functions associated with HPV16 viral load in CIS and SCC, respectively. We further show that viral loads were unexpectedly low early in the SCC disease process.

Impact: HPV16 viral load appears highly complex which may limit its use in cervical screening. *Cancer Epidemiol Biomarkers Prev*; 22(1); 150–8. ©2012 AACR.

Introduction

Numerous studies have shown an association between high copy numbers of human papillomavirus 16 DNA (HPV16 viral load), persistent infection (1–4), and malignant transformation of the cervical epithelium (2, 5–14). However, some studies have found a weak association (15) or even lower HPV16 viral load

in high-grade compared with low-grade cervical intraepithelial neoplasia (CIN), possibly because integration of the viral genome disrupts viral production (16). Others have predicted limitations of using viral load for screening purposes due to large within-disease variation in copy numbers and substantial overlap between women with and without disease (17–20). Coinfection with other HPV types and surrounding low-grade lesions may also affect the viral load measurement (21, 22). Hence, the value of viral load technology for large-scale screening is uncertain.

Viral load might, however, increase screening specificity in triaging HPV16-positive women (23) and could serve as a marker/predictor of cervical carcinogenesis (24). Currently, the field is hampered by the limited prospective data on viral load and risk for invasive disease that has only rarely been included as an outcome in prospective studies (7). Cross-sectional studies have included such invasive disease in large numbers (10, 13) but do not allow risk predictions because the exposure and outcome are not separated in time. Our aim was to prospectively study the significance of HPV16 viral load in *in situ* and invasive cancer, which should increase our knowledge about the natural history of HPV infections, and the potential use of HPV16 viral load in screening.

Authors' Affiliations: ¹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet; ²Swedish eScience Center (SeRC), Stockholm; ³Department of Laboratory Medicine, Karolinska University Hospital, Huddinge; ⁴The WHO HPV LabNet Global Reference Laboratory, Department of Laboratory Medicine, Lund University, Malmö; ⁵Division of Clinical Cancer Epidemiology, Department of Oncology, Sahlgrenska University Hospital, Gothenburg, Sweden; ⁶Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts; and ⁷Institute for Molecular Medicine Finland, Helsinki University, Helsinki, Finland

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

K. Sundström and A. Ploner contributed equally to this work.

Corresponding Author: Karin Sundström, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O. Box 281, S-171 77 Stockholm, Sweden. Phone: 460-852-482-364; Fax: 460-831-4975; E-mail: karin.sundstrom@ki.se

doi: 10.1158/1055-9965.EPI-12-0953-T

©2012 American Association for Cancer Research.

Materials and Methods

Participants

As the mid-1970s, all Swedish women have been invited for cytological screening with Papanicolaou (Pap) smears every 3 to 5 years (25). Records containing all information from the screening are computerized in the National Cervical Screening Register (NCSR) and virtually all smears have been stored (26). The highly reliable National Cancer Register (NCR) records all new diagnoses of both cancer *in situ* and invasive cancer since 1958 (27). These databases can be linked through the individually unique Swedish personal ID numbers.

Our source population comprised all women (1,459,258) who participated in cervical screening within 1 of 10 county laboratories in 6 Swedish counties sometime during 1969 to 2002. We used the NCSR to identify a cohort of 1,431,724 women whose first registered smear (defining cohort entry) was classified as cytologically normal (Pap = 1). This cohort was then linked to the NCR to identify all women with a first histologically confirmed diagnosis of cancer *in situ* (CIS), or invasive squamous cell carcinoma (SCC) after cohort entry. A diagnosis of CIS in the NCR translates internationally to a diagnosis of CIN grade 3 (CIN3). As the incidence of CIS is higher than the incidence of SCC (28, 29), we included only a random sample of CIS to achieve case groups of more similar size.

To be eligible as a control, a woman could not have had a histologically confirmed diagnosis of CIS/SCC at the time of diagnosis of the case. Among all eligible control women, 1 control woman—matched on county laboratory, date of cohort entry (+/− 3 months), and age at first normal smear (+/− 1 year)—was randomly selected for each CIS and SCC case. All available smears from the case patient and the control woman taken before the date of diagnosis of the case were retrieved from biobank archives.

To verify the original histologic diagnosis, all available histologic specimens from the case women were rereviewed by a senior pathologist. This was done to conduct sensitivity analyses using different classifications of disease. Because of the multiplicity of pathology laboratories involved (in total 38 such), the overall proportion of missing histologies in this study was 18% for SCC and 21% for CIS. Given our nation-wide approach and study time spanning several decades, we hold this to be acceptable given that our previous study (6) had a missing rate of 11%, in just 1 county laboratory. Some laboratories were contracted to resection new cases from old paraffin-embedded tissue, as the original slides had been disposed of. In several of these instances, due to sparsity of material to resection, our pathologist was not able to review all original tissue blocks representatively. It is, therefore, not certain whether all the cases that were reclassified as lower grade than originally truly are of lower grade, or that our pathologist was not able to review them to the same precision as the original diagnostician. In combination with the fact that all cases in the study were originally histopathologically verified; the issue of reduced power due to missing histologic rereview and partially unclear

validity of the reclassification means that we hold presentation of both the full and the reclassified data to be the most informative.

Smear analyses

Each smear was recoded and relabeled to ensure blinding of case–control status; samples belonging to the same matched case–control pair were included in the same analysis batch at the same calendar time. DNA extraction was conducted by validated methods (30). All smears were analyzed for the presence of 7 low-risk HPV types (HPV 6, 7, 11, 42, 43, 70, and 90), and 16 high-risk HPV types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, 73, and 82).

The PCR amplification of a consensus region using GP5+/6+ primers (31) was followed by HPV type detection through a PCR-enzyme immunosorbent assay (EIA) and reverse dot blot hybridization procedure (32) or detection of biotinylated HPV amplicons by a multiplex fluorescent bead-based assay (33). Presence of amplifiable DNA in samples was determined by PCR-EIA or real-time PCR for the housekeeping β -globin gene. All analyses were conducted for the current study; no smears/HPV data were reused from our previous publication on another sample of CIS (6).

In HPV16-positive samples, we further quantified the viral load, measured as an absolute number of viral copies of the E7 gene per microliter, using the Taqman real-time quantitative PCR method (34). HPV analyses were conducted in the WHO HPV LabNet Global Reference Laboratory, Malmö, Sweden. The analyzing laboratory was blinded to the identity of the samples.

Because quantitative β -globin was not available for the entire material, we did not normalize the viral load variable to viral copies per cell or amount of DNA. In samples where quantitative β -globin was available, we found high correlation between absolute and normalized viral load values ($r = 0.86$ for CIS and $r = 0.80$ for SCC). Furthermore, it can be noted that the use of a count of cells as a denominator also has limitations. This, as it is not known how many cells in the sample/smear were actually infected with HPV16, or with what copy number, and the number of viral copies per cell, therefore, represents an average. Previous research has also indicated that absolute level of viral load is associated with risk for CIN on its own without adjustment for sample cellularity (35), irrespective of the use of a housekeeping gene as a denominator (36) and that risk estimates for CIN2/CIN3 were hardly affected by adjusting for such a gene (37).

Initially, 6,567 smears were available for statistical analyses. We subsequently excluded 150 smears with negative β -globin value, 337 smears that had not undergone viral load analysis and 415 smears from 182 incomplete case–control pairs (because either the case or the control had no retrievable Pap smears). Hence, 5,665 smears from 621 complete case–control pairs of CIS and 457 case–control pairs of SCC remained.

Statistical methods

In descriptive case–case comparison analyses, we evaluated multiple time scales for their relevance for the dynamics of viral load. Because controls had few HPV16-positive smears, we modeled viral loads among case women only. We fitted linear models for the logarithmized (base 10) viral loads with time before diagnosis as primary time scale; and then systematically added individual and combined age, period, and cohort (APC) effects to the basic model. These effects were evaluated both for their statistical significance and size. These preliminary models indicated weak or no effect of birth cohort and age at smear on viral load in either case group (Supplementary Table S1). However, we did find a significant effect of calendar period of smear for CIS, where viral loads decreased systematically over time: the highest levels were found during the 1980s, with levels at the end of the 1990s almost an order of magnitude lower. Sensitivity analysis indicated that this period effect affected risk estimates (described below) for CIS, but not for SCC; consequently, only viral loads for CIS were adjusted for period of smear (Supplementary Figs. S1 and S2, and Supplementary Tables S2 and S3).

Independently of these adjustments, we also found that viral loads for cases and controls varied systematically within the period of HPV analyses, with later measurements tending to be lower during the analysis period (2005–2010). Consequently, all viral loads were adjusted for calendar date of analysis.

Risk estimation and viral load imputation

We used conditional logistic regression to study the association between HPV16 viral load and risk of CIS and SCC at 1 to 10 years before diagnosis, using HPV16-negative women as reference. A practical consideration when modeling the risk associated with viral load cross-sectionally over time was the irregular pattern in which smears were available due to our study being nested in a screening program with variable participant compliance. This meant that few matched case–control pairs originally had measurements in the same intervals. As a consequence, only few matched pairs would contribute to the estimation of each cross-sectional OR, leading to wide confidence intervals and problems with separation. To address this issue, we implemented a within-person imputation scheme for each individual, as done previously by our group (6). For each subject and each year, we imputed HPV16 measurements from all observations within a fixed window (± 5 years), where observations closer in time received larger weights (see Supplementary Data).

For the final results, we conducted 100 imputation runs for all individuals, resulting in 100 imputed data sets. The final estimate of the OR at each time point t was calculated as the average of the estimates from the imputed data sets. The corresponding variance was calculated as the sum of the average of the within-imputation variance and the between-imputation variance, according to Rubin's for-

mula (38). Alternative values of 4 and 6 years for the time window w had little effect on the results (data not shown). The ORs were interpreted as relative risks, as cervical cancer is a rare disease outcome, and the controls were sampled from the underlying population at risk.

We used a marginal mean approach to modeling longitudinal effects in the viral loads of case women and the risk of cervical cancer associated with HPV16 viral load. Individual trajectories of viral load in women with multiple smears can vary substantially (data not shown), which is why we opted for the per-subject imputation scheme outlined above for estimating risk. Our results are consequently valid and interpretable on an average population level.

We further conducted sensitivity analyses using histologically reconfirmed cases only, and analyses restricted to the pre-1995 period, as that period was covered by our former study (6). Calculations were done using R version 2.14.0 (39).

The study was approved by the Karolinska Institutet Ethics Review Board, Stockholm, Sweden, which also determined that informed consent from the participants was not required.

Results

Participants

Median age at diagnosis was 33 years in CIS, and 41 years in SCC, with a median follow-up time from the first smear of 6.1 years for CIS, and 7.7 years for SCC (Table 1). Except for the last year before diagnosis, the number of smears was similar between cases and controls. The percentage of smears positive for any HPV type was higher among CIS cases than SCC cases, but the percentage of HPV16-positive smears was the same at 33% to 34%. Among controls, the percentage of smears positive for any HPV type was 12% to 15% and for HPV16 3% to 4% (Table 1).

Eighty-one percent of the SCC cases were, upon histologic rereview, again classified as SCC, 13% reclassified as CIS and 6% reclassified as other (e.g., anaplastic cancer and adenocarcinoma). Of the CIS cases, 62% were again classified as CIN3, 27% reclassified as CIN2, 8% reclassified as CIN1, and 3% reclassified as SCC.

Case–case comparison

Average HPV16 viral load levels increased continuously in HPV16-positive smears from CIS cases from 10 years before diagnosis (Fig. 1), with the strongest surge during the last year before diagnosis. In contrast, among SCC cases, viral load levels remained more stable from 10 to 4 years before diagnosis, but then gradually increased, especially during the last 2 years before diagnosis (Fig. 1).

For CIS cases, the level of viral load differed between calendar periods, so that women who were followed for 10 years in the earliest period (1980–1990) had the highest viral load at diagnosis and women who were followed for 10 years until they developed CIS in the latest period

Table 1. Characteristics of the study participants. Counts and continuous variables are reported as median (5–95 percentile)

	CIS		SCC	
	Cases	Controls	Cases	Controls
Women	621	621	457	457
Age at diagnosis (in years) ^a	33 (23–51)	33 (23–51)	41 (28–67)	41 (28–67)
Age at entry (in years)	25 (18–45)	26 (18–45)	33 (20–59)	33 (20–59)
Time in study (in years)	6.1 (0.4–15.4)	6.2 (0.8–15.6)	7.7 (0.3–17.8)	8.0 (1.2–18.8)
Total number of smears	1,740	1,489	1,265	1,177
Number of smears/subject	2 (1–13)	2 (1–14)	2 (1–18)	2 (1–15)
Number of smears [0,1] years before diagnosis	520	167	256	109
(1,5] years	559	589	342	377
(5,10] years	444	494	382	403
(10,15] years	171	182	201	196
(15,19] years	27	30	53	50
HPV-positive smears (%)	65	15	55	12
HPV16-positive smears (%)	34	4	33	3

^aFor controls, age at time of diagnosis of the matched case woman.

(1995–2005) had the lowest. A similar, yet not statistically significant, trend was visible among SCC cases (Fig. 1).

We further observed that the groups of women later diagnosed with CIS or SCC on average initially had low viral loads of 5 to 10 copies/microliter, especially in the later calendar periods from 10 years to around 4 years before diagnosis and in SCC (Fig. 1). The late surge resulted in average viral loads of around 100 copies/microliter in both case groups at diagnosis in the latest periods.

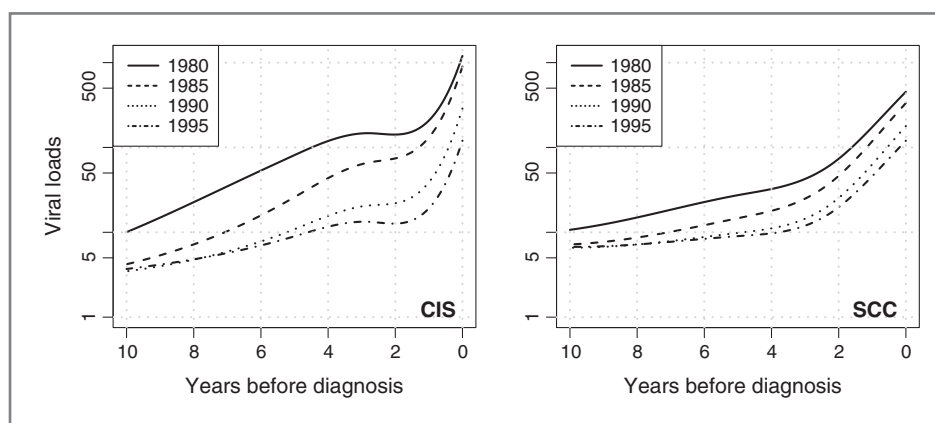
We investigated extensively whether these period effects in viral load could be explained by technical aspects in our study. As noted above, an assay drift was detected, but this was related to calendar date of analysis, not calendar period of smear. Also, it could not explain the stronger calendar period effect on viral load in CIS than in SCC. Time trends in histologic revision did not explain the calendar period effect either. Furthermore, no time trend

in HPV16 positivity over the study period was detected (data not shown).

Case-control analysis

In the imputation model for CIS, the medium and the highest tertiles of viral load aligned similarly in terms of risk throughout 10 years before diagnosis (Fig. 2, upper left). These medium/high tertile categories were, therefore, collapsed in the final CIS imputation model to increase precision (Fig. 2, lower left). (For category values of viral load and relative risks associated with individual tertiles, see Supplementary Tables S4–S6.) In CIS, being HPV16-positive and having the lowest tertile of viral load was associated with a constant 10-fold excess risk compared with HPV16-negative women throughout 10 years before diagnosis. However, a medium/high viral load was associated with an excess risk of CIS from around

Figure 1. Average HPV16 viral loads (in copies of HPV16/ μ L) in HPV16-positive smears from case women with CIS (left) and SCC (right). Viral loads are modeled via natural spline functions on a log scale (base10) over 10 years before diagnosis. Viral loads are further stratified according to 5-year calendar period of the starting point. For example, the solid black line indicates the viral load trajectory for the group of women who start contributing smears in 1980 and were diagnosed in 1990.



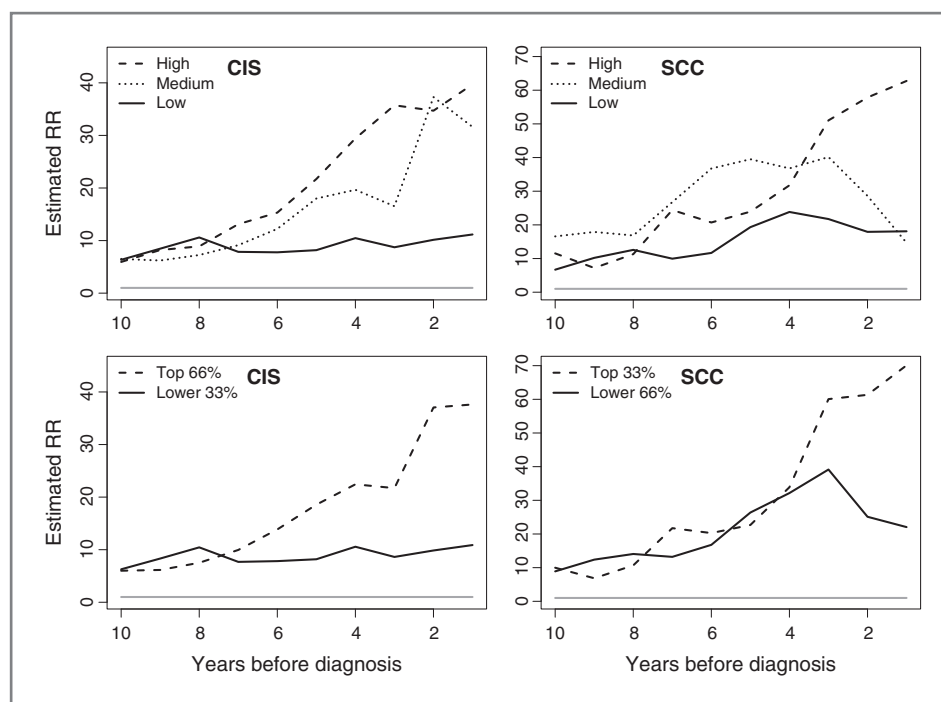


Figure 2. Imputed relative risks for squamous cervical cancer *in situ* (CIS, left) and squamous cell carcinoma (SCC, right) according to level of HPV16 viral load compared with HPV16-negative subjects over 10 years before diagnosis. Upper panels display separated tertiles and lower panels display low versus collapsed medium/high (CIS) and collapsed low/medium versus high (SCC), respectively. The gray horizontal line at RR = 1 is the reference level risk of HPV16-negative.

7 years before diagnosis, with a relative risk of around 40 during the last year before diagnosis (Table 2). These risk estimates were robust in sensitivity analyses using data from pre-1995 only (Supplementary Fig. S3) and histologically reconfirmed CIS cases only, respectively, although with lower precision due to a reduction in power (Supplementary Table S7).

Overall HPV16-positivity (regardless of viral load level) conferred a risk profile for CIS similar to the low viral load category compared with HPV16-negative women (Table 2), except for the 2 to 3 last years before diagnosis where the risk assumed a more intermediate position between the 2 viral load categories. High-risk HPV positivity conferred much the same risk profile as

Table 2. Left section: imputed ORs for CIS according to level of HPV16 viral load compared with HPV16-negative women over 10 years before diagnosis; right section: imputed ORs for CIS according to HPV16-positivity and high-risk HPV positivity over 10 years before diagnosis^a

HPV16 viral load and risk of cervical cancer <i>in situ</i> (CIS)				HPV16-positivity and HR-HPV-positivity and risk of CIS		
Year before diagnosis	HPV16 negative (reference)	Low HPV16 viral load (OR, 95% CI)	Medium/high HPV16 viral load (OR, 95% CI)	Year before diagnosis	HPV16 positive Ref: HPV16-negative (OR, 95% CI)	High-risk HPV positive Ref: High-risk HPV-negative (OR, 95% CI)
1	1.0	11 (3–36)	38 (11–130)	1	17 (9–33)	25 (14–48)
2	1.0	10 (3–29)	37 (11–127)	2	15 (8–28)	19 (11–33)
3	1.0	9 (3–22)	22 (8–59)	3	12 (7–20)	14 (9–22)
4	1.0	11 (4–29)	22 (8–65)	4	13 (8–24)	14 (9–23)
5	1.0	8 (3–20)	19 (7–48)	5	11 (7–20)	13 (8–20)
6	1.0	8 (3–18)	14 (5–37)	6	9 (6–16)	9 (6–14)
7	1.0	8 (3–17)	10 (4–24)	7	8 (5–14)	8 (6–13)
8	1.0	10 (4–29)	8 (3–19)	8	9 (5–15)	8 (5–13)
9	1.0	8 (3–21)	6 (2–17)	9	8 (4–14)	7 (4–11)
10	1.0	6 (3–16)	6 (2–17)	10	7 (4–13)	6 (4–9)

^aNote the different reference levels for the different analyses.

Table 3. Left section: imputed ORs for SCC according to level of HPV16 viral load compared with HPV16-negative women over 10 years before diagnosis; right section: imputed ORs for SCC according to HPV16-positivity and high-risk HPV positivity over 10 years before diagnosis^a

HPV16 viral load and risk of squamous cell carcinoma (SCC)				HPV16-positivity and HR-HPV-positivity and risk of SCC		
Year before diagnosis	HPV16 negative (reference)	Low/medium HPV16 viral load (OR, 95% CI)	High HPV16 viral load (OR, 95% CI)	Year before diagnosis	HPV16 positive Ref: HPV16-negative (OR, 95% CI)	High-risk HPV positive Ref: High-risk HPV-negative (OR, 95% CI)
1	1.0	22 (6–81)	70 (8–628)	1	20 (8–50)	29 (12–72)
2	1.0	25 (6–104)	61 (7–547)	2	18 (8–43)	22 (10–49)
3	1.0	39 (8–202)	60 (6–580)	3	20 (8–50)	20 (9–43)
4	1.0	32 (7–159)	34 (4–271)	4	17 (7–39)	17 (9–36)
5	1.0	26 (6–112)	23 (4–119)	5	17 (8–38)	16 (8–32)
6	1.0	17 (6–50)	20 (3–139)	6	12 (6–26)	11 (6–20)
7	1.0	13 (5–35)	22 (3–154)	7	9 (5–17)	8 (5–14)
8	1.0	14 (5–38)	11 (2–54)	8	9 (4–17)	7 (4–12)
9	1.0	12 (5–33)	7 (2–27)	9	8 (4–17)	9 (5–15)
10	1.0	9 (3–23)	10 (2–57)	10	8 (4–15)	7 (4–13)

^aNote the different reference levels for the different analyses.

that of HPV16-positivity, except slightly higher in the last year before diagnosis.

In SCC, the low and medium tertiles were similar and thus were collapsed (Fig. 2, upper and lower right). Between 10 to 3 years before diagnosis of the case, both the resulting categories (low/medium vs. high viral load) were similar in terms of risk for SCC. This common risk increased from about 10 at 10 years before diagnosis, ending at a relative risk of about 35 at 4 years before diagnosis compared with an HPV16-negative woman (Fig. 2, lower right). Therefore, for the larger part of the follow-up period the risk for invasive cancer appeared mainly associated with HPV16-positivity, of any viral load level. However, from around 4 years before diagnosis the highest category of viral load was associated with a clear separation in risk. Finally, during the last year before diagnosis the highest tertile of viral load conferred an almost 70-fold risk, compared with a 20-fold risk associated with low/medium viral load (Table 3). In sensitivity analyses, risk estimates for the pre-1995 period were similar to the original results (Supplementary Fig. S4). When using histologically reconfirmed SCC cases only, the risk separation between the low/medium and highest viral load tertile at 4 years before diagnosis was also robust although somewhat less dramatic, ultimately yielding a 25-fold increased risk associated with low/medium viral load and a 40-fold increased risk associated with high viral load (Supplementary Table S8).

In SCC, HPV16-positivity and high-risk HPV positivity aligned with the risk conferred by the low/medium viral load category (Table 3 and Supplementary Table S6).

Discussion

In this comprehensive study, we describe HPV16 viral load in squamous cervical carcinogenesis. We previously showed the significance of consistent high viral load of HPV16 in cervical CIS (6). Using a similarly stringent study design and an entirely new data collection, we present a new large sample of CIS and in addition a large sample of women with invasive cancer (SCC), allowing unique comparative analyses.

We show that women with CIS had a qualitatively different increase in viral load over time compared with women with SCC. The lack of overlap between the 2 curves was an unexpected finding and illustrates the potential complexity of viral load kinetics. This difference could well be a manifestation of viral integration because such integration may lower virion production but still promote progression to invasive neoplasia (16). The plateau phase among the SCC cases may also relate to the fact that we included only SCC with a first cytologically normal smear at entry into the cohort, a strategy chosen to enable a more homogenous case group and study of the progression from a normal smear to diagnosis of cancer. These may represent a group of lesions, which for a number of years could lie relatively indolent and escape detection. Following a surge in viral load the last 2 years before diagnosis, they may finally have been detected. In contrast, the larger or more productive lesions—either of which can result in higher viral loads—could be more likely to be caught earlier before giving rise to SCC, thus more often presenting as cases of CIS.

In our previous study (6), a strong association between HPV16 viral load and risk of CIS was seen already many years before diagnosis, although it should be noted that

the measure used for viral load was indirect (threshold cycle number). We here confirm this finding for absolute numbers of viral load: medium/high viral load is a risk factor for CIS in at least a 7-year perspective, in line with recent findings that viral load of HPV16 predicts progression of infection (40). We also found an increased risk for SCC associated with high viral load, although the threshold for risk separation was different in the 2 diagnoses. Furthermore, having a lower viral load was also associated with an excess risk for both CIS and SCC compared with being HPV16-negative.

Our finding of low viral loads early in the SCC disease process may have additional implications. While the first priority of cervical screening programs is the detection of precancerous cervical lesions, a secondary goal is early detection and downstaging of invasive disease. Our data indicate that less sensitive HPV testing than our PCR could have failed to detect low-grade infections that still correlated with invasive disease in the long-term perspective. This is in line with a previous study where HPV tests taken more than 6 years before diagnosis were not predictive of invasive cervical cancer (41). Therefore, although the current recommendation is to use relatively high analytical thresholds to locate only HPV infections at clear risk for prevalent disease (36, 42), our data support the notion that to determine the optimal duration of protection for future disease, a lower threshold may be considered. The invasive cancer cases in our study can be considered interval cancers occurring despite some measure of regular screening. For this group of women with a first registered cytologically normal smear, highly sensitive thresholds may be of importance in the long term. However, this gain in sensitivity should be weighed against a subsequent loss in specificity in the larger population of women—an issue of import in large-scale HPV-based screening. It should also be noted that our viral load measure originated from archival smears and has been subject to technical adjustments as described.

The strengths of this study include the stringent epidemiological design, nation-wide approach, comprehensive HPV testing and statistical analysis, and availability of histologic rereview. Although some cross-sectional studies, and 1 prospective study, have included moderate to large numbers of invasive cancers (7, 10, 13) there is, to our knowledge, no other prospective study which can present viral load results for large samples of CIS and SCC within the same population-based cohort. Indeed, ongoing cohort studies will likely not yield such data, as women presenting with precancerous lesions will always receive treatment before developing invasive disease.

Limitations of our study include the great variation in numbers of smears. Also, we sampled only women participating in screening. However, in the study of screening tests, this is unavoidable and does not reflect selection bias. The shown calendar period effects on

viral load are enigmatic and may represent a combination of technical and biological issues. Possibly, cervical screening may have become more effective at detecting CIS of a smaller size—resulting in a lower viral load at diagnosis. As lesion size was not known in this material, we could not directly evaluate that factor. All the same, adjustment for the calendar period effect takes period into account and importantly thus allows for valid interpretations of the results. Although a proportion of our cases were histologically reclassified upon rereview, this is to be expected given the known variation in such diagnoses between pathologists (29). Also, our risk estimates were robust in sensitivity analyses indicating that this was not of central importance. Our imputation model relies on several assumptions to estimate cross-sectional ORs per year and the resulting confidence intervals are overlapping due to the uncertainty added. Nevertheless, it yields a qualitative result that is well in line with previous literature for CIN3/CIS. We, therefore, consider this approach to be reasonable also for SCC risk prediction.

Integration status—that is, whether HPV DNA is found in episomal form or integrated in the host cell genome—is typically calculated through measuring the number of E2 gene copies (a gene frequently disrupted during integration) and comparing the ratio of this to the number of E6 or E7 gene copies (genes that remain intact during integration; ref. 43). Because our assay targeted only E7, we did not assess integration which is a limitation given our findings in invasive cancer. However, the exact timing of viral integration during malignant transformation of cervical epithelium is disputed (11), and in a recent cross-sectional study, HPV16 viral load was significantly associated with cancer irrespective of E2 status (13). Others have also concluded that viral load alone can be used, although a combination with E2/E6 ratio adds information (44). Hence, not including integration status in our risk prediction models may not be a major limitation.

In conclusion, we show that high viral load of HPV16 clearly predicts risk for both *in situ* and invasive squamous cervical cancer. However, the risk functions differed in this study, suggesting that viral load quantification cannot easily strengthen the positive effect of HPV-based screening. We further show that copy numbers can be low in early stages of invasive cancer, which may be worth considering when weighing sensitivity against specificity in HPV-based cervical screening.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: L. Arnheim Dahlström, J. Dillner, H.-O. Adami, N. Ylitalo, P. Sparén

Development of methodology: A. Ploner, J. Palmgren, J. Dillner, H.-O. Adami

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): K. Sundström, L. Arnheim Dahlström, J. Dillner, P. Sparén

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): K. Sundström, A. Ploner, J. Palmgren, J. Dillner, H.-O. Adami, P. Sparén

Writing, review, and/or revision of the manuscript: K. Sundström, A. Ploner, L. Arnheim Dahlström, J. Dillner, H.-O. Adami, N. Ylitalo, P. Sparén

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): K. Sundström, A. Ploner, H.-O. Adami

Study supervision: L. Arnheim Dahlström, J. Dillner, H.-O. Adami, N. Ylitalo, P. Sparén

Acknowledgments

The authors thank Ms. Ninoa Malki for database management; Dr. Anders Lindgren for histopathologic review; Ms. Carina Eklund, Ms. Christina Cavala, Ms. Aline Marshall, and Ms. Kia Sjölin for laboratory

analyses; Dr. Anthony Gunnell for data collection; and Ms. Kristina Glimsjö for secretarial support.

Grant Support

This work was supported by the National Cancer Institute at the NIH (grant numbers 1R01CA093378-01A1 and 5R01CA111720-03). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the NIH. The funders had no role in study design, data collection and analysis, preparation of the manuscript, or decision to publish the manuscript.

Received August 20, 2012; revised October 23, 2012; accepted November 5, 2012; published OnlineFirst November 15, 2012.

References

- Munoz N, Hernandez-Suarez G, Mendez F, Molano M, Posso H, Moreno V, et al. Persistence of HPV infection and risk of high-grade cervical intraepithelial neoplasia in a cohort of Colombian women. *Br J Cancer* 2009;100:1184–90.
- Schlecht NF, Trevisan A, Duarte-Franco E, Rohan TE, Ferenczy A, Villa LL, et al. Viral load as a predictor of the risk of cervical intraepithelial neoplasia. *Int J Cancer* 2003;103:519–24.
- Dalstein V, Riethmuller D, Pretet JL, Le Bail Carval K, Sautiere JL, Carbillat JP, et al. Persistence and load of high-risk HPV are predictors for development of high-grade cervical lesions: a longitudinal French cohort study. *Int J Cancer* 2003;106:396–403.
- Maucort-Boulch D, Plummer M, Castle PE, Demuth F, Safaeian M, Wheeler CM, et al. Predictors of human papillomavirus persistence among women with equivocal or mildly abnormal cytology. *Int J Cancer* 2010;126:684–91.
- Josefsson AM, Magnusson PK, Ylitalo N, Sorensen P, Qvarforth-Tubbin P, Andersen PK, et al. Viral load of human papilloma virus 16 as a determinant for development of cervical carcinoma *in situ*: a nested case-control study. *Lancet* 2000;355:2189–93.
- Ylitalo N, Sorensen P, Josefsson AM, Magnusson PK, Andersen PK, Ponten J, et al. Consistent high viral load of human papillomavirus 16 and risk of cervical carcinoma *in situ*: a nested case-control study. *Lancet* 2000;355:2194–8.
- Moberg M, Gustavsson I, Wilander E, Gyllensten U. High viral loads of human papillomavirus predict risk of invasive cervical carcinoma. *Br J Cancer* 2005;92:891–4.
- Monnier-Benoit S, Dalstein V, Riethmuller D, Lalaoui N, Mouglin C, Pretet JL. Dynamics of HPV16 DNA load reflect the natural history of cervical HPV-associated lesions. *J Clin Virol* 2006;35:270–7.
- Snijders PJ, Hogewoning CJ, Hesselink AT, Berkhof J, Voorhorst FJ, Bleeker MC, et al. Determination of viral load thresholds in cervical scrapings to rule out CIN 3 in HPV 16, 18, 31 and 33-positive women with normal cytology. *Int J Cancer* 2006;119:1102–7.
- Wu Y, Chen Y, Li L, Yu G, Zhang Y, He Y. Associations of high-risk HPV types and viral load with cervical cancer in China. *J Clin Virol* 2006;35:264–9.
- Cricca M, Morselli-Labate AM, Venturoli S, Ambretti S, Gentilomi GA, Gallinella G, et al. Viral DNA load, physical status and E2/E6 ratio as markers to grade HPV16 positive women for high-grade cervical lesions. *Gynecol Oncol* 2007;106:549–57.
- Gravitt PE, Kovacic MB, Herrero R, Schiffman M, Bratti C, Hildesheim A, et al. High load for most high risk human papillomavirus genotypes is associated with prevalent cervical cancer precursors but only HPV16 load predicts the development of incident disease. *Int J Cancer* 2007;121:2787–93.
- Das D, Bhattacharjee B, Sen S, Mukhopadhyay I, Sengupta S. Association of viral load with HPV16 positive cervical cancer pathogenesis: causal relevance in isolates harboring intact viral E2 gene. *Virology* 2010;402:197–202.
- Rodriguez AC, Schiffman M, Herrero R, Hildesheim A, Bratti C, Sherman ME, et al. Longitudinal study of human papillomavirus persistence and cervical intraepithelial neoplasia grade 2/3: critical role of duration of infection. *J Natl Cancer Inst* 2010;102:315–24.
- Castle PE, Schiffman M, Scott DR, Sherman ME, Glass AG, Rush BB, et al. Semiquantitative human papillomavirus type 16 viral load and the prospective risk of cervical precancer and cancer. *Cancer Epidemiol Biomarkers Prev* 2005;14:1311–4.
- Woodman CB, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. *Nat Rev Cancer* 2007;7:11–22.
- Xi LF, Kiviat NB, Galloway DA, Zhou XH, Ho J, Koutsky LA. Effect of cervical cytologic status on the association between human papillomavirus type 16 DNA load and the risk of cervical intraepithelial neoplasia grade 3. *J Infect Dis* 2008;198:324–31.
- Hesselink AT, Berkhof J, Heideman DA, Bulkman NW, van Tellingen JE, Meijer CJ, et al. High-risk human papillomavirus DNA load in a population-based cervical screening cohort in relation to the detection of high-grade cervical intraepithelial neoplasia and cervical cancer. *Int J Cancer* 2009;124:381–6.
- Boulet GA, Benoy IH, Depuydt CE, Horvath CA, Aerts M, Hens N, et al. Human papillomavirus 16 load and E2/E6 ratio in HPV16-positive women: biomarkers for cervical intraepithelial neoplasia >or = 2 in a liquid-based cytology setting? *Cancer Epidemiol Biomarkers Prev* 2009;18:2992–9.
- Constantinou-Williams C, Collins SI, Roberts S, Young LS, Woodman CB, Murray PG. Is human papillomavirus viral load a clinically useful predictive marker? A longitudinal study. *Cancer Epidemiol Biomarkers Prev* 2010;19:832–7.
- Xi LF, Edelstein ZR, Meyers C, Ho J, Chene SL, Schiffman M. Human papillomavirus types 16 and 18 DNA load in relation to coexistence of other types, particularly those in the same species. *Cancer Epidemiol Biomarkers Prev* 2009;18:2507–12.
- Sherman ME, Wang SS, Wheeler CM, Rich L, Gravitt PE, Tarone R, et al. Determinants of human papillomavirus load among women with histological cervical intraepithelial neoplasia 3: dominant impact of surrounding low-grade lesions. *Cancer Epidemiol Biomarkers Prev* 2003;12:1038–44.
- Gravitt PE, Coutlee F, Iftner T, Sellors JW, Quint WG, Wheeler CM. New technologies in cervical cancer screening. *Vaccine* 2008;26 Suppl 10:K42–52.
- Franco EL, Coutlee F. Prognostic value of measuring load of human papillomavirus DNA in cervical samples: an elusive target. *J Natl Cancer Inst* 2009;101:131–3.
- Gustafsson L, Sparen P, Gustafsson M, Wilander E, Bergstrom R, Adami HO. Efficiency of organised and opportunistic cytological screening for cancer *in situ* of the cervix. *Br J Cancer* 1995;72:498–505.
- Sparen P. Cervical cancer screening in Sweden in 2006. Report from the National Cervical Cancer Screening Register. Stockholm: Karolinska Institutet; 2007.
- Barlow L, Westergren K, Holmberg L, Talback M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. *Acta Oncol* 2009;48:27–33.
- The Swedish National Board of Health and Welfare. Cancer Incidence in Sweden 2007. Official Statistics of Sweden; 2008.

29. Bergstrom R, Adami HO, Gustafsson L, Ponten J, Sparen P. Detection of preinvasive cancer of the cervix and the subsequent reduction in invasive cancer. *J Natl Cancer Inst* 1993;85:1050–7.
30. Chua KL, Hjerpe A. Polymerase chain reaction analysis of human papillomavirus in archival cervical cytologic smears. *Anal Quant Cytol Histol* 1995;17:221–9.
31. de Roda Husman AM, Walboomers JM, van den Brule AJ, Meijer CJ, Snijders PJ. The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. *J Gen Virol* 1995;76:1057–62.
32. Soderlund-Strand A, Rymark P, Andersson P, Dillner J, Dillner L. Comparison between the Hybrid Capture II test and a PCR-based human papillomavirus detection method for diagnosis and posttreatment follow-up of cervical intraepithelial neoplasia. *J Clin Microbiol* 2005;43:3260–6.
33. Soderlund-Strand A, Carlsson J, Dillner J. Modified general primer PCR system for sensitive detection of multiple types of oncogenic human papillomavirus. *J Clin Microbiol* 2009;47:541–6.
34. Moberg M, Gustavsson I, Gyllensten U. Real-time PCR-based system for simultaneous quantification of human papillomavirus types associated with high risk of cervical cancer. *J Clin Microbiol* 2003;41:3221–8.
35. Gravitt PE, Burk RD, Lorincz A, Herrero R, Hildesheim A, Sherman ME, et al. A comparison between real-time polymerase chain reaction and hybrid capture 2 for human papillomavirus DNA quantitation. *Cancer Epidemiol Biomarkers Prevent* 2003;12:477–84.
36. Snijders PJ, van den Brule AJ, Meijer CJ. The clinical relevance of human papillomavirus testing: relationship between analytical and clinical sensitivity. *J Pathol* 2003;201:1–6.
37. van Duin M, Snijders PJ, Schrijnemakers HF, Voorhorst FJ, Rozendaal L, Nobbenhuis MA, et al. Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance. *Int J Cancer* 2002;98:590–5.
38. Rubin DB. Multiple imputation for non-response in surveys. New York: John Wiley & Sons; 1987.
39. Development Core Team R. A language and environment for statistical computing. R Foundation for Statistical Computing. [Accessed November 27, 2012]. Available at <http://www.R-project.org>; 2008.
40. Xi LF, Hughes JP, Castle PE, Edelstein ZR, Wang C, Galloway DA, et al. Viral load in the natural history of human papillomavirus type 16 infection: a nested case-control study. *J Infect Dis* 2011;203:1425–33.
41. Wallin KL, Wiklund F, Angstrom T, Bergman F, Stendahl U, Wadell G, et al. Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. *N Engl J Med* 1999;341:1633–8.
42. Kinney W, Stoler MH, Castle PE. Special commentary: patient safety and the next generation of HPV DNA tests. *Am J Clin Pathol* 2010;134:193–9.
43. Andersson S, Safari H, Mints M, Lewensohn-Fuchs I, Gyllensten U, Johansson B. Type distribution, viral load and integration status of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). *Br J Cancer* 2005;92:2195–200.
44. Fontaine J, Hankins C, Mayrand MH, Lefevre J, Money D, Gagnon S, et al. High levels of HPV-16 DNA are associated with high-grade cervical lesions in women at risk or infected with HIV. *AIDS* 2005;19:785–94.