Gynecologic Oncology 154 (2019) 345–353

Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

Association of age and viral factors with high-risk HPV persistence: A retrospective follow-up study

Wending Li [a,1], Yifan Meng [b,l,1], Yi Wang [a,1], Xiaodong Cheng [c,1], Chen Wang [d,1], Songshu Xiao [d], Xiaofei Zhang [c], Zaixing Deng [e], Mengjun Hu [f], Pingrong Shen [g], Shengfeng Xu [h], Chenglin Fu [i], Wen Jiang [j], Bing Wu [k], Kezhen Li [b,l], Gang Chen [b,l], Juncheng Wei [b,l], Ling Xi [b,l], Junbo Hu [l], Ding Ma [b,l], Min Xue [d,***], Xing Xie [c,**], Peng Wu [b,l,*]

[a] School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[b] Department of Gynecologic Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[c] Department of Gynecologic Oncology, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China
[d] Department of Obstetrics and Gynecology, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China
[e] Department of Obstetrics and Gynecology, Huzhou Maternity and Child Care Center, Huzhou, Zhejiang, China
[f] Department of Obstetrics and Gynecology, Zhuji People's Hospital, Shaoxing, Zhejiang, China
[g] Department of Obstetrics and Gynecology, Ninghai Maternity and Child Care Center, Ningbo, Zhejiang, China
[h] Department of Obstetrics and Gynecology, Jiaxing Maternity and Child Care Center, Jiaxing, Zhejiang, China
[i] Department of Obstetrics and Gynecology, Taizhou First People's Hospital, Taizhou, Zhejiang, China
[j] Department of Obstetrics and Gynecology, Zhoushan Maternity and Child Care Center, Zhoushan, Zhejiang, China
[k] Department of Obstetrics and Gynecology, Zhuji Maternity and Child Care Center, Shaoxing, Zhejiang, China
[l] The Key Laboratory of Cancer Invasion and Metastasis of the Ministry of Education of China, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

HIGHLIGHTS

- We found a time-dependent association of age and viral factors with high-risk HPV clearance.
- Older age inhibited high-risk HPV clearance only after 400 days of infection.
- Co-infection promotes high-risk HPV clearance in the start, but the effect attenuated and reversed as infection persisted.
- Recurrent same-type infection cleared slower than the previous one.

ARTICLE INFO

Article history:
Received 17 January 2019
Received in revised form 6 May 2019
Accepted 28 May 2019
Available online 24 June 2019

Keywords:
Human papillomavirus
Persistence
Risk factor
Retrospective

ABSTRACT

Objective. Cervical HR-HPV persistence is the main risk factor for cervical cancer. We aimed to investigate the association of age and viral factors with HR-HPV persistence.

Methods. From 2010 to 2017, 343,128 women underwent 390,411 tests performed by the Cervista HR-HPV assay (Data C3) and 157,123 women underwent 206,505 tests performed by the GenoArray HR-HPV assay (Data G14) in nine medical centers located in central and eastern China. We combined the test results and identified 9234 HPV-specific baseline-negative records for time-to-event analyses. The study endpoint event was defined as clearance of type/group-specific HPV. Therefore, hazard ratio (HR) < 1 indicated a higher risk of HPV persistence, which is contrary to the common meaning of HR.

Results. The median persistence time was 375 and 541.5 days for Data C3 and Data G14, respectively. For every 5-year increase in age, a 15% (95% confidence interval [CI], 11%–19%) decrease in the clearance rate was observed only after 400 days of infection. For each additional co-infected HPV, the HR was 1.80 (95% CI, 1.63–1.97) on infection initiation but decreased by 22% (95% CI, 18%–26%) every 100 days. The HR of infection

Abbreviations: CI, confidence interval; CIN, cervical intraepithelial lesion; HPV, human papillomavirus; HR, hazard ratio; HR-HPV, high-risk human papillomavirus.

* Correspondence to: P. Wu, Department of Gynecologic Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China.
** Correspondence to: X. Xie, Department of Gynecologic Oncology, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China.
*** Correspondence to: M. Xue: Department of Obstetrics and Gynecology, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China.
E-mail addresses: xueminxy3@163.com (M. Xue), xiex@zju.edu.cn (X. Xie), pengwu8626@tjh.tjmu.edu.cn (P. Wu).

[1] Wending Li, Yifan Meng, Yi Wang, Xiaodong Cheng, Chen Wang contributed equally in this work.

https://doi.org/10.1016/j.ygyno.2019.05.026
0090-8258/

recurrence was 0.48 (95% CI, 0.32–0.72). The findings were consistent across different populations and test methods and were robust in sensitivity analysis.

Conclusions. We found a time-dependent association of age and viral factors with HPV clearance. Older age reduced HPV clearance only after 400 days of infection. Co-infection promoted HPV clearance in the beginning, but the effect attenuated and reversed as infection persisted. Recurrent same-type infection cleared slower than the previous one.

1. Introduction

High-risk human papillomavirus (HR-HPV) infection is known to be responsible for some vulvar and vaginal cancers, anal cancers, head and neck cancers, and most cervical cancers [1]. Numerous epidemiologic studies and clinical trials have demonstrated cervical HR-HPV persistence as the main risk factor for high-grade cervical intraepithelial neoplasia (CIN) and cervical cancer [2,3]. With an estimated 570,000 cases and 311,000 deaths worldwide in 2018, cervical cancer ranks fourth for both incidence and mortality in women [4]. Additionally, primary HR-HPV testing has been recommended for cervical cancer screening according to updated clinical guidelines released by the American Society for Colposcopy and Cervical Pathology (ASCCP) and EUropean Research Organization on Genital Infection and Neoplasia (EUROGIN) [5,6]. Presently, HR-HPV-positive patients are recommended to undergo 12-month follow-ups and repeated HPV tests regardless of age, HPV type, infection times, and co-infection status differences. Therefore, it is important to understand the natural history and risk factors of HR-HPV persistence for developing personalized screening strategies as attempted in many cohort studies [7–28]. However, most studies restricted their analyses to HPV-positive women at study entry, resulting in an underestimation of the duration of HR-HPV persistence because the infection time before recruitment was unknown [15–28]. In other studies that had restricted duration calculation in baseline-negative women, the effective sample sizes (the number of infection episodes) were small owing to the prospective design [7–14]. This dilemma has led to a general inconsistency of findings of the traits and risk factors of HR-HPV persistence. Furthermore, it remained unclear how co-infection and recurrent same-type HPV infection would influence the persistence time. In this study, we used real-world data from nine medical centers in China to identify eligible baseline-negative patients (in women testing HPV-negative at enrollment, the proportion with a newly detected HPV infection at any subsequent screening visit) and conducted an exploratory large-scale retrospective study to examine the natural HPV infection cycle traits.

2. Methods

2.1. Study population, HPV tests, and study factors

The reporting of this study adheres to the REporting of studies Conducted using Observational Routinely collected health Data (RECORD) checklist [29]. From November 17, 2010, to April 28, 2017, we recorded a total of 596,916 tests routinely performed in the Departments of Obstetrics and Gynecology of nine large medical centers located in seven cities of China (Supplementary Table S1). The tests were performed for women who (i) had a history of sexual activity, (ii) were not using vaginal medications at the time of enrollment, (iii) had never been vaccinated for HPV, and (iv) consented to HPV testing. No other exclusion criteria were imposed. Among these tests, 390,411 tests from the cities of Wuhan, Hangzhou, Huzhou, Ningbo, Shaoxing, and Zhoushan were performed using the Cervista HPV HR Assay (Hologic Inc., Bedford, MA, USA) and 206,505 tests from Changsha City were performed using the HPV GenoArray Diagnostic test kit (Hybribio Ltd., Guangzhou, Guangdong, China). The Cervista categorizes 14 HR-HPV types into three groups and provides positive results in the form of HPV groups, which include A5/A6 (HPV51, 56, and 66), A7 (HPV18, 39, 45, 59, and 68), and A9 (HPV16, 31, 33, 35, 52, and 58) [30]. The GenoArray can determine 15 HR-HPV types and six low-risk HPV types [31]. However, we included the same 14 HR-HPV types as those in the Cervista to maintain consistency of data analysis. Although the only FDA-approved assay is for HPV 16 or 18 genotyping, Cervista HPV HR and GenoArray assays are widely used in co-testing cervical screening program in China together with the physician's assessment of cytology history. The two methods have proved to be highly sensitive and specific for the grouping or genotyping of HPV and provide excellent predictive values [30,31]. Procedures to avoid cross-contamination are as follows: (i) Use nuclease-free, sterile disposable aerosol barrier pipette tips for each addition and transfer. (ii) Wear gloves when setting up the test. (iii) Make sure that the pipette tips touch only the solution being dispensed. More information regarding the study population, specimen collection, and test methods is given in Supplementary Table S2 and Supplementary Methods.

2.2. Age and viral factors

Age was calculated as years between the date of birth and the date of the first positive test in the infection episode. HPV groups included A5/A6 (HPV51, 56, and 66), A7 (HPV18, 39, 45, 59, and 68), and A9 (HPV16, 31, 33, 35, 52, and 58), as given by the test method of Cervista. HPV types included all the 14 HR-HPV types covered by A5/A6, A7, and A9. Co-infection is coded as 0 for no co-infection of other group/type, 1 for single co-infection, 2 for two co-infections, and so forth. Number of total hospital visits included both positive and negative tests. Recurrent infection is coded as 1 and 2 for the first and the second observed infection episode, respectively. As it was impossible to know the number of times the infection of a certain HR-HPV type/group had occurred and cleared before the current infection, all analyses concerning recurrent infection were restricted to women who had two same-type/same-group infection episodes recorded in our data.

2.3. Data processing and datasets

HPV raw data generated with the Cervista and GenoArray methods were processed separately (Fig. 1). Briefly, we combined multiple test results from the same patient and identified 26,501 women who had never been admitted as inpatients and had been taking more than one HPV test with at least one of them positive. We further divided them and obtained 33,897 HPV-specific records (i.e., each record represented the infection episode of a specific HPV type/group). Among them, 5233 and 4001 baseline-negative records showed the use of Cervista and GenoArray methods and thus were named as Data C3 and Data G14 accordingly (3 and 14 refers to 3 HPV groups and 14 HPV types that these two datasets represented, respectively). Henceforth, all primary analyses were conducted and validated across Data C3 and Data G14. However, as HPV group-based Data C3 and HPV type-based Data G14 could not be combined into one dataset directly, we separated the 14 HPV types of Data G14 into 3 groups as per the Cervista method and combined it with Data C3. This combined dataset, named as Data Combine, was used only in Cox regression analyses to obtain overall hazard ratios (HRs) when Data C3 and Data G14 yielded similar results.

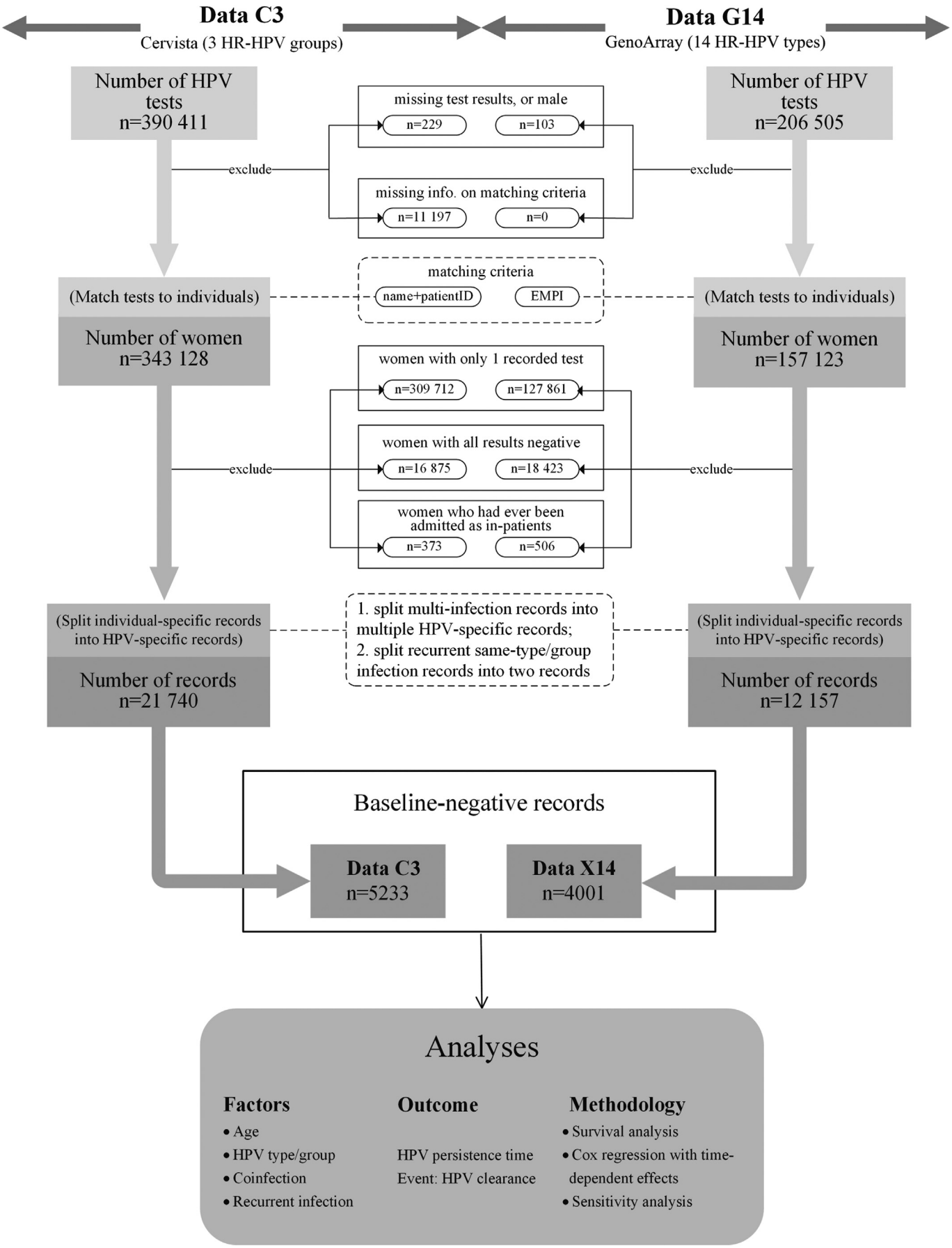

Fig. 1. Flowchart of data processing and analyses.

2.4. Ascertainment of HPV infection time

A detailed graphical illustration of infection time calculation is shown in Supplementary Fig. S1. Two methods were used to calculate the HPV infection time. The exact-time method, as in conventional time-to-event analysis, estimates the exact infection time as the time span between the midpoint of negative-to-positive conversion dates and the midpoint of the following positive-to-negative conversion dates or the last positive date, whereas the interval-censored method does not estimate the exact infection time but provides the interval that the actual infection time must fall within. The latter approach, although suffers from reduced power and limited statistical utilization, can compensate for the risk of misestimating the real HPV infection time with the midpoint approach of the exact-time method (e.g., should the distribution of two actual conversion dates skew disproportionally to one side), and it was thus used to test the robustness of our findings. All analyses were performed using the exact-time method unless otherwise specified.

2.5. Survival analysis and Cox regression

Survival analysis was performed separately for Data C3 and Data G14. We constructed and compared HPV persistence curves by different variables to gain an overview of their effects and possible interactions with infection time.

To validate the primary findings from survival analysis and estimate effect sizes, we constructed Cox models for Data C3, Data G14, and Data Combine. The endpoint event was defined as the clearance of a type/group-specific HPV. We used "clearance rate" to refer to the "hazard" of HPV clearance. In our study, HR < 1 indicated higher risk of HPV persistence, which is contrary to the common meaning of HR. Moreover, we used "time-varying variables" to refer to variables that may change values at different HPV tests during infection (e.g., co-infection status may vary from one test to another during one HPV-specific infection episode. Therefore, the whole infection episode should be divided into multiple segments according to the co-infection situation of each visit and different co-infection values should be allocated to each segment accordingly). We also used "time-dependent variables" to refer to variables with significant interactions with time (i.e., variable effect changes as infection persists) [32,33]. Traditional Cox regression assumes proportional hazard over time (i.e., constant hazard ratio); however, there are cases where this assumption fails to apply. By including a time-dependent variable, an interaction term between time and the covariate under study, the HR is allowed to change with time. The HR of a covariate can then be divided into two parts: the coefficient of the covariate term, which can be interpreted as the HR at time 0, and the coefficient of the time-interaction term, which indicates the relative change in HR for every one-unit increase in time. More information about the rationale and interpretation of time-dependent variable is given in Supplementary Methods.

Age was regarded as a time-dependent continuous variable; the time effect was binary and had a cutoff time point of 400 days. (This cutoff was inferred from observation of survival curves and validated in subsequent sensitivity analysis.) Co-infection was treated as a time-varying, time-dependent variable (the time dependency was also based on observation of prior survival analysis and validated in Cox regression). We also included the number of total hospital visits and its interaction term with time into the model. However, if the time effect in the time-interaction term was expressed in days, then the relative change in risk, even if highly significant, would be too small to be of any numerical or clinical significance for a one-day increase in time (e.g., an HR of 0.999 with 95% CI of 0.998 to 0.999). Therefore, we divided the time effect in the time-interaction terms of co-infection and number of hospital visits by 100, and thus, the unit of time inflated to 100 days. To investigate the effect of infection recurrence, we restricted all relevant analyses to a subgroup of women with two same-type/same-group HPV infection episodes recorded (a total of 75 and 59 women were identified for Data C3 and Data G14, respectively).

2.6. Sensitivity analyses

Considering possible interdependence among multiple type/group-specific records from the same person, we additionally constructed dependence-adjusted Cox models using the modified sandwich estimator method. This method does not require prior assumptions of the nature of dependence and can estimate weighted averages of all coefficients in one SAS PHREG procedure [32]. We also tested the robustness of our findings in the previously mentioned interval-censored records (which covers exactly the same records but provides infection time intervals instead of time point estimation) using the ICPHREG procedure, a similar SAS procedure specifically designed for analyzing interval-censored data. We compared the HRs of all variables across seven cities and pooled them together using a random-effects model. We also used restricted cubic splines regression to test for possible nonlinearity of age in Cox regression modeling.

Analyses were performed using SAS software (version 9.4, SAS Institute, Cary, USA). *P* values of <0.05 were considered to indicate statistical significance.

3. Results

3.1. Basic characteristics and comparability of Data C3 and Data G14

Supplementary Table S2 presents the basic characteristics of women included in Data C3 and Data G14. A total of 5233 and 4001 baseline-negative records were extracted for Data C3 and Data G14, respectively. For Data C3 and Data G14, the mean ages of women were 41.8 and 43.7 years, the median persistent infection times were 375.0 and 541.5 days, and the clearance rates were 483 and 342 per 1000 infections per year, respectively. Supplementary Fig. S2 displays the HPV persistence curve of two datasets. Data G14 presented a lower clearance probability than Data C3 at any comparable time.

We evaluated possible causes of the curve difference between Data C3 and Data G14 and identified four testable factors that may cause such difference: age, method of HPV grouping/typing, method of identifying the same study subject, and possible calculation error introduced by the exact-time method. After analysis, we found that all four factors were likely not the main cause of the difference (see Supplementary Material B). However, the difference among city-specific curves was apparent and significant from one another ($P < 0.001$, Supplementary Fig. S3), suggesting that the cause might be related to the disparity among different study populations (e.g., differential distribution of HPV types or socioeconomic factors).

3.2. Overview of factor effects

For both Data C3 and Data G14, age had no effect until approximately day 400, when a turning point was observed for older age groups (especially the "60-high" group; Fig. 2A, B).

HR-HPV persistence curves varied substantially across different types/groups. In Data C3, the A9 group showed lower clearance probability than the A5/A6 and A7 groups (Fig. 3A). HPV16, 33, 39, 52, and 58 were more likely to persist, whereas HPV45, 56, and 59 cleared faster (Fig. 3B).

We used records from patients with two same-type/same-group infection episodes to compare survival curves of different infection times (Fig. 4A, B). Recurrent infection was significantly harder to clear than the previous one in both Data C3 ($P = 0.003$) and Data G14 ($P = 0.025$).

Co-infections were more likely to clear in Data C3 ($P < 0.001$) but not in Data G14 ($P = 0.967$, Fig. 4C, D). However, for Data G14, the curve of the ever-co-infected was below the curve of the never-co-infected before day 400 ($P = 0.002$ when confining the comparison to

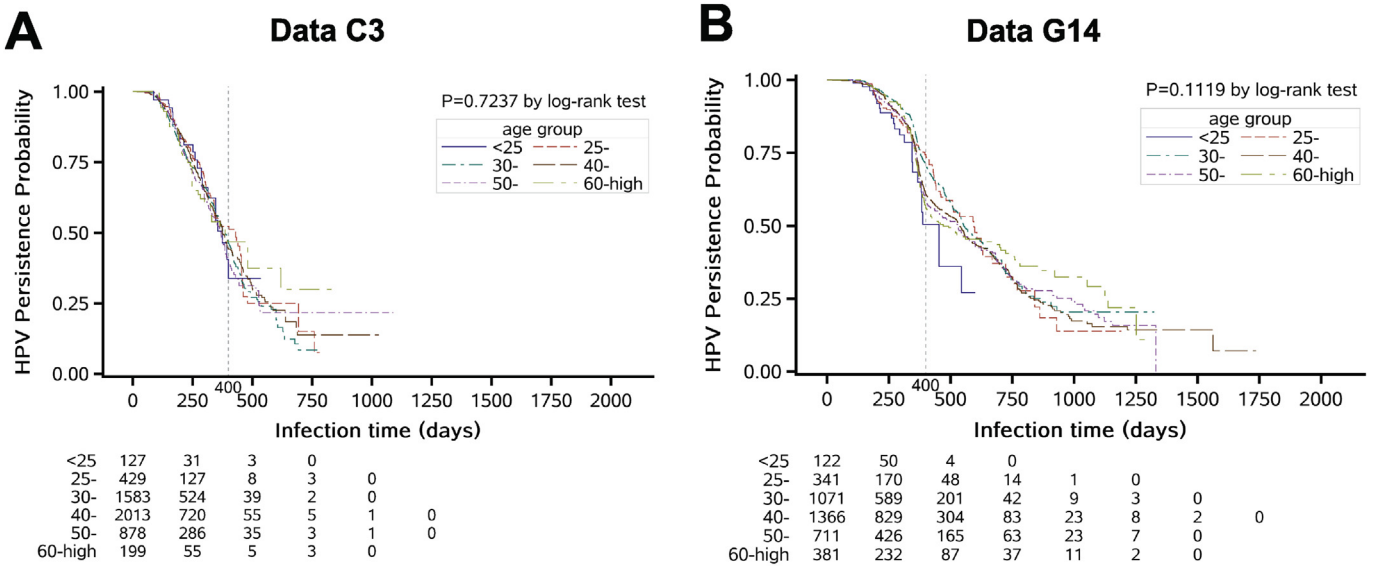

Fig. 2. Age-specific HR-HPV infection persistence curves, by data source. Data C3 was defined as the dataset of records that used the Cervista test method (with three high-risk HPV groups). Data G14 was defined as the original dataset of records that used the GenoArray test method (with 14 high-risk HPV types).

<400 days), but the curves converged and reversed afterward (Fig. 4D). Likewise, the curves of the ever-co-infected and the never-co-infected of Data C3 also tended to converge after day 400 (Fig. 4C). This observation suggests possible time dependence of co-infection.

3.3. Cox regression

Using the exact-time Cox regression with variables invariant and independent of time (the conventional Cox proportional hazards model), the HRs of age, number of visits, and recurrent infection were similar across Data C3 and Data G14 (Supplementary Fig. S4). The effect of co-infection reversed from Data C3 to Data G14, indicating possible interaction with time. Results of time-dependent Cox regression are presented in Table 1 and Supplementary Fig. S5. The effect of age was significant only after 400 days, when every 5-year increase in age was associated with a 15% decrease in the clearance rate (95% confidence interval [CI], 11%–19%). Compared with the A9 group, the HRs of the A5/A6 and A7 groups were 1.43 (95% CI, 1.30–1.58) and 1.22 (95% CI, 1.11–1.34), respectively, for the Data Combine dataset. HPV16, 33, 39, 52, and 58 were less likely to clear, whereas HPV45, 56, and 59 cleared faster (all presented HPV types had >20% change in HR compared with HPV51; Supplementary Table S3). Co-infection was significantly associated with an increased clearance rate (HR of HPV clearance, 1.80; 95% CI, 1.63–1.97), and the effect seemed to attenuate as infection persisted (for every 100 days, the HR decreased by 22%; 95% CI, 18%–26%). Therefore, co-infection only promoted HR-HPV clearance during an early period and reversed to clearance-inhibiting after 233 days of infection. The number of hospital visits and its time-dependent term were significant

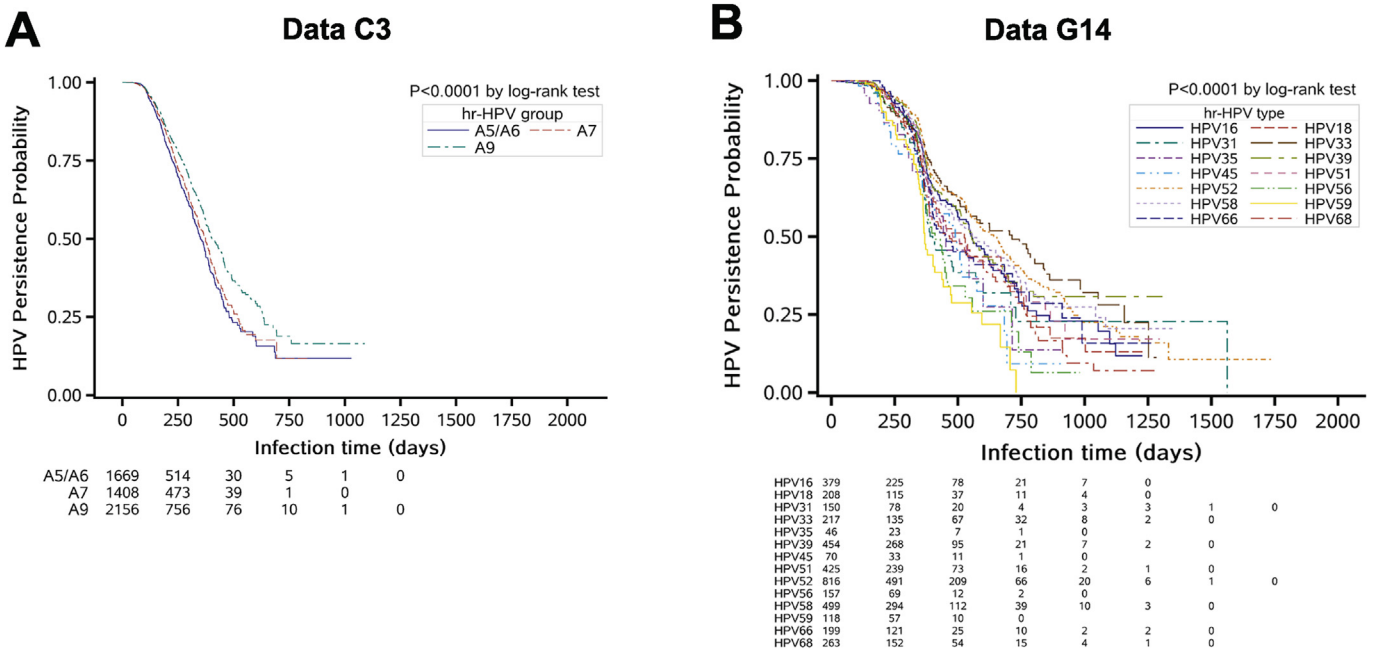

Fig. 3. Type/group-specific HR-HPV infection persistence curves, by data source. Data C3 was defined as the dataset of records that used the Cervista test method (with three high-risk HPV groups). Data G14 was defined as the original dataset of records that used the GenoArray test method (with 14 high-risk HPV types).

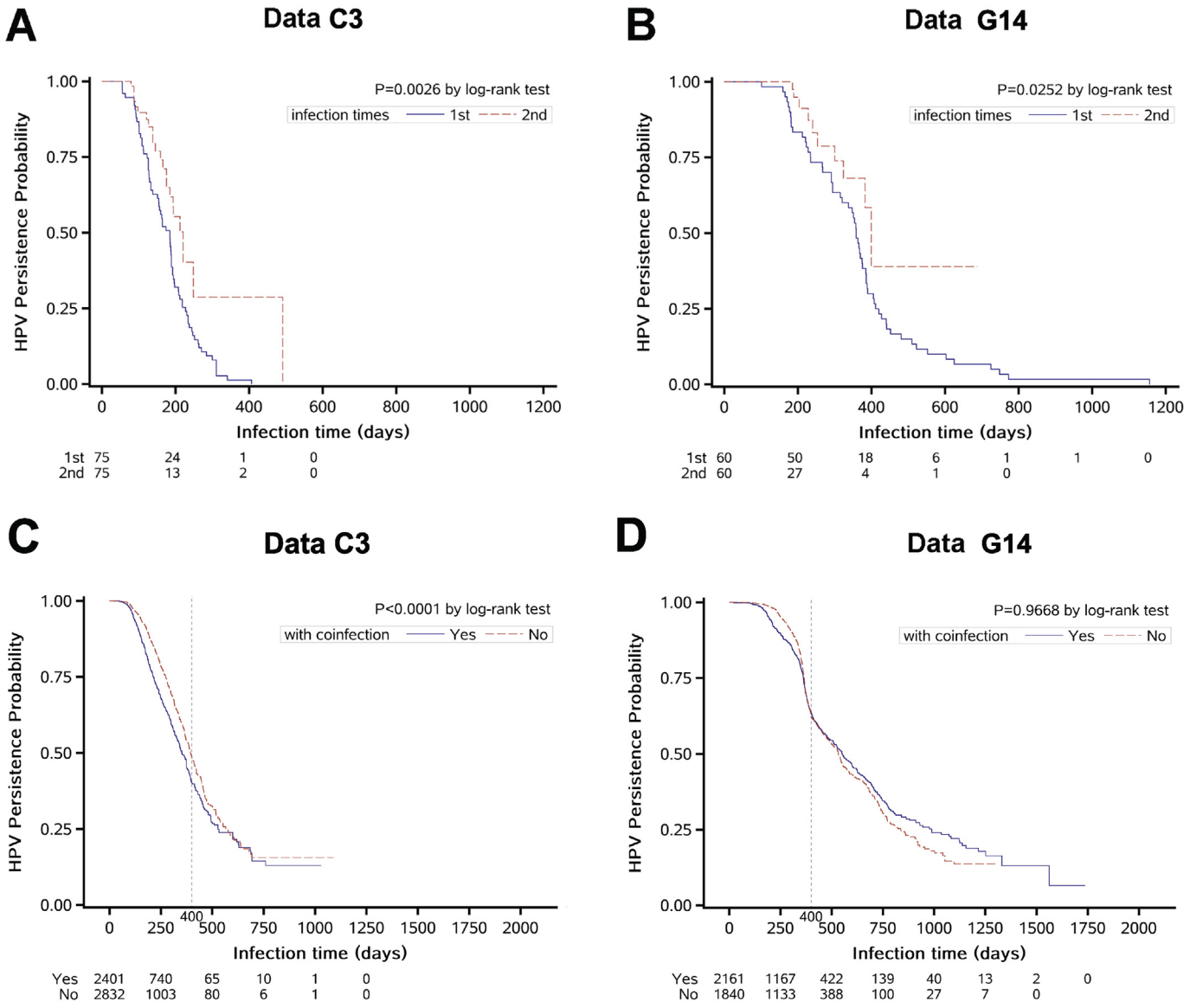

Fig. 4. Infection times-specific and co-infection status-specific HR-HPV infection persistence curves, by data source. Survival analysis was restricted to women who had recurrent same-type/same-group infection episodes recorded in infection times-specific analysis. A "yes" for co-infection indicated that there existed at least one test during which the patient was infected with more than one HPV type/group.

and robust across Data C3 and Data G14. Among patients with recurrent same-type/same-group infection, each recurrence decreased the clearance rate by 52% (95% CI, 28%–68%).

3.4. Sensitivity analyses

The HRs and significance remained largely unchanged when we re-entered infection time as interval-censored data (Supplementary Fig. S4) or adjusted for data interdependence (Supplementary Fig. S5). To test the robustness of the association of age and viral factors with HR-HPV persistence, we conducted a meta-analysis with data grouped by city (Fig. 5). As shown in Fig. 5, the HRs remained generally consistent across different cities and were close to HRs presented in Table 1. Therefore, Supplementary Fig. S2 and Fig. 5 together showed that although HR-HPV persistence curves varied across datasets and cities, the factor effects were robust.

Because the choice of the 400-day time point as cutoff for age's time interaction term was based on observation of unadjusted survival analysis and was thus subject to certain arbitrariness, we constructed the same Cox models as those given in Table 1 with different cutoff time points for age (Supplementary Fig. S6). An apparent HR differentiation between pre-cutoff and post-cutoff age curves was observed at day 400 for all datasets, suggesting that our previous choice of a 400-day cutoff was adequate. We also used restricted cubic spline regression to construct the same models as those given in Table 1 and found no evidence to reject our modeling assumption that a linear relationship existed between age (as a continuous variable) and logarithm hazard (P for nonlinearity was >0.15 for both pre- and post-400 day age effects of all datasets; details not shown).

4. Discussion

As persistent HR-HPV infection is the key step in the development of cervical cancer, it is well recognized that extensive screening programs are important to prevent cervical cancer, especially for unvaccinated women in developing countries. Therefore, annual HPV re-examination has been widely recommended for HR-HPV-positive population. However, our findings suggest that women with different HPV infection profiles might undergo different infection experiences, and a personalized screening strategy might be a more efficient and economical choice. In

Table 1
Hazard ratio (95% CI) for high-risk HPV clearance and *P* values in time-dependent Cox regression, by dataset.

	Data C3 $N = 5233$		Data G14 $N = 4001$		Data Combine $N = 8561$	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age (per 5 yr)						
Infection <400 d	0.99 (0.96 to 1.02)	0.463	1.00 (0.97 to 1.03)	0.253	1.00 (0.97 to 1.02)	0.667
Infection ≥400 d	0.85 (0.78 to 0.94)	<0.001	0.88 (0.84 to 0.92)	<0.001	0.85 (0.81 to 0.89)	<0.001
HPV group						
A5/A6	1.42 (1.25 to 1.60)	<0.001	NA	NA	1.43 (1.30 to 1.58)	<0.001
A7	1.17 (1.03 to 1.33)	0.018	NA	NA	1.22 (1.11 to 1.34)	<0.001
Co-infection at t_0	1.43 (1.29 to 1.59)	<0.001	1.49 (1.31 to 1.70)	<0.001	1.80 (1.63 to 1.97)	<0.001
Co-infection × t (per 100 d)	0.90 (0.85 to 0.95)	<0.001	0.87 (0.84 to 0.92)	<0.001	0.78 (0.74 to 0.82)	<0.001
N_visit at t_0	1.84 (1.73 to 1.96)	<0.001	1.86 (1.71 to 2.01)	<0.001	1.80 (1.72 to 1.88)	<0.001
N_visit × t (per 100 d)	0.93 (0.91 to 0.96)	<0.001	0.93 (0.91 to 0.95)	<0.001	0.94 (0.93 to 0.96)	<0.001
Recurrent infection	0.44 (0.26 to 0.74)	0.002	0.42 (0.20 to 0.88)	0.021	0.48 (0.32 to 0.72)	<0.001

Abbreviations: N, number of records, each representing one infection episode of one HPV type/group of one woman; HR, hazard ratio; CI, confidence interval; N_visit, number of hospital visits; NA, not applicable. Data C3 was defined as the dataset of records that used the Cervista test method (with three high-risk HPV groups). Data G14 was defined as the original dataset of records that used the GenoArray test method (with 14 high-risk HPV types). Data Combine was defined as the combined dataset of Data C3 and Data G14 with 14 HPV types grouped into three as in Data C3. Data C3 and Data Combine were additionally adjusted for city. The HRs of co-infection and number of hospital visits each comprised two parts: the HR at time 0 (t_0) and the relative change in HR with every 100-day increase in time (the product term with t, in 100 d). The hazard ratio (95% CI) and P value of each HPV type for Data G14 are shown in Supplementary Table 3. The estimation of recurrent infection was restricted to a subgroup of patients who had two same-type/same-group infection episodes (Data C, $N = 150$; Data G14, $N = 118$; Data Combine, $N = 268$). The reference level of the HPV Group was A9.

this study, we found that age, HPV group/type, co-infection status, number of hospital visits, and recurrent infection were associated with HR-HPV persistence in different ways. The findings were robust across different datasets and cities and remained largely unchanged when adjusted for data interdependence or interval censoring. To our knowledge, our sample size was the largest among studies assessing HPV persistence and its risk factors. This advantage has facilitated in-depth analyses that were barely conceivable in previous studies (e.g., time-dependent analysis, interval-censored analysis, dependence adjustment, and analysis of recurrent infection) [7–26].

Older age increased the risk of HR-HPV persistence after 400 days of infection. Seven out of eight studies defining a baseline-positive infection showed a significant association or trend of increased risk of HPV persistence with older age [15,20–23,25,26]; the exception had only 1 year of follow-up [24]. However, for four studies defining a negative-to-positive infection initiation [8–11], only one study showed a significant association [11]. This observation was understandable, as our study suggested that the overall (time-independent) effect of age was marginal and might lose significance with increased standard error and short follow-up time. The delayed effect of age on HR-HPV persistence may be attributed to high viral load, high proportion of viral integration, and high risk of co-infection caused by physiological and immunological disorders that concur with aging [34]. Furthermore, epigenetic reactivation of previously silenced integration sites may also play a role [35].

In line with other studies, we confirmed that the A9 group was the most persistent HR-HPV group [8,9,14,18,21]. We also found that HPV16, 33, 39, 52, and 58 were more likely to persist, whereas HPV45, 56, and 59 cleared faster. Although no consensus had been reached regarding the most persistent HPV types (most favored were HPV33 and HPV16), our results were generally consistent with those of previous studies [7–9,13,15,18,19,21,22]. The discrepancy in the persistence of

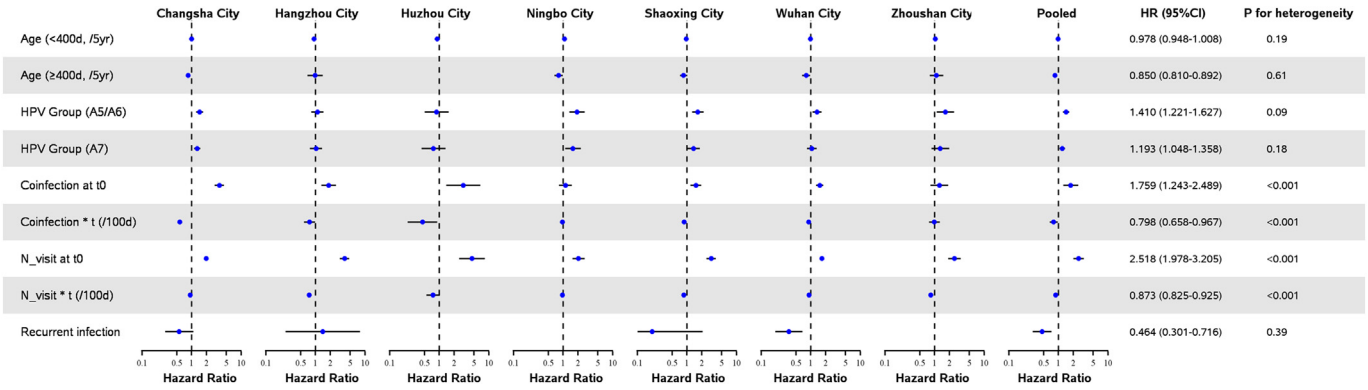

Fig. 5. City-specific and pooled hazard ratio of high-risk HPV clearance in time-dependent Cox regression Abbreviations: HR, hazard ratio; CI, confidence interval; N_visit, number of hospital visits. Hazard ratios from different cities were pooled using the random-effects model. Some results were not shown owing to insufficient number. The HRs of co-infection and number of hospital visits each comprised two parts: the HR at time 0 (t0) and the relative change in HR with every 100-day increase in time (the product term with t, in 100 d). The estimation of recurrent infection was restricted to a subgroup of patients who had two same-type/same-group infection episodes. The reference level of the HPV Group was A9. Horizontal lines represent 95% confidence intervals.

different HR-HPV types suggests that type-specific carcinogenesis may be different due to HR-HPV genetic structure and integration probability [36].

We found that, co-infected HR-HPV infections cleared faster early than no co-infection, but the effect attenuated with time and reversed after 8 months of infection. Many previous studies treated co-infection only as a time-invariant variable, which was inaccurate, as the number of co-infections may vary throughout one infection episode. Of nine related studies, a significant increase in co-infection-associated risk of persistence was found in four [8,9,11,24] and the other five found no significance [10,19,21,22,26]. One possible explanation for our finding is that the host's immune system might be more activated when first encountering multiple infections but rendered less capable to clear HPV with time. The effect of viral interaction may also change as infection persists [37].

For women who had contracted and cleared HPV infection in the past, the existence and the magnitude of naturally acquired protection against homologous HPV reinfection are still in dispute [38]. To our knowledge, our study provided the first epidemiological evidence that recurrent same-type HR-HPV infection is less likely to clear. As the recorded first infection most likely did not occur for the first time in one's infection history, such effect should be cumulative. The previous HPV infection might have altered the host's cellular DNA or immune system, thus debilitating the host's ability to clear HPV by herself. HPV detectability, sexual behavior factors, physiological changes, and immunological factors may also contribute. However, caution needs to be taken as the "reinfection" might actually be reactivation of a latent HPV infection, or deposition of HPV from a partner, and not an infection per se. Every newly detected "infection" has the same issue, but it requires special attention when interpreting the re-detection of the same-type HR-HPV.

Women who had visited hospital more frequently were more likely to experience HPV clearance; yet, the effect slightly attenuated as infection persisted. Hospital visits were included in the model to reflect two aspects that were not measured directly in our study, that is, comorbidities or ill health, which may drive HPV-infected women to visit hospital more frequently, and self-awareness and concerns for the current HPV infection. As we excluded patients who had ever been admitted as inpatients, the unrecorded comorbidities, if existed, should be some mild general gynecological diseases. In addition, the data partly supported the self-awareness assumption: if the woman was more concerned with her infection and visited hospital for recheck more frequently, she may take other measures (including avoiding sex or cleaning the vulva) to help with HPV clearance, and these measures did help from the beginning. However, because of either a lack of full adoption of these measures or other comorbidities, this protective effect attenuates as infection persists.

The median HR-HPV persistence times were 375 and 541.5 days for Data C3 and Data G14, respectively, and ranged from 224 to 601 days in existing publications with a negative-to-positive definition of HR-HPV infection initiation [8–14,39]. These findings suggest that although it might be possible to determine the effect of risk factors on HR-HPV clearance, the crude HR-HPV persistence curves could behave differently across different populations.

Our study has several limitations. Because of its retrospective design, we were unable to assess other risk factors such as lifestyle and history of disease, which might have explained underlying differences among populations. Although we excluded women who had been admitted as in-patients, we lacked direct information to distinguish women with cervical morbidities. However, only 2.7% (4288/157123) of the initial population of Data G14 had a recorded CIN or cervical cancer, which suggested our findings should not have been substantially confounded by cancer-related cervical morbidity. Possible selection bias was another weakness. However, attempts were made to improve the generalizability of our findings by including diverse study populations and two testing methods and by using the number of hospital visits to reflect patients' concerns about HPV infection.

5. Conclusion

We identified older age, HPV group/type, co-infection, and recurrence of infection as risk factors for HR-HPV persistence. The effects of age and co-infection varied with time, and 400 days was determined to be an important turning point. These findings suggested that, instead of setting a general cutoff, personalized follow-up intervals should be recommended (e.g., based on age, HPV type, co-infection status, and infection times). Nonetheless, our study was large but exploratory, and further corroborations from prospective studies are needed.

Author contributions

Peng Wu, Xing Xie, Min Xue, and Wending Li conceived and designed the study; Wending Li and Yi Wang performed the statistical analysis; Yifan Meng, Yi Wang, and Wending Li drafted the article; Min Xue, Songshu Xiao, Chen Wang, Xiaofei Zhang, Zaixing Deng, Mengjun Hu, Pingrong Shen, Shengfeng Xu, Chenglin Fu, Wen Jiang, Bing Wu, Kezhen Li, Ling Xi, Ding Ma, Xing Xie, Junbo Hu, Gang Chen, Juncheng Wei, and Xiaodong Cheng contributed to the data collection and quality control; Peng Wu, Xing Xie, and Min Xue made critical revisions to the manuscript. All authors revised and commented on the article and approved the final version before submission.

Ethical approval

This study was approved by the Ethical Committee of Tongji Hospital of Tongji Medical College at Huazhong University of Science and Technology (TJ-IRB20171001). Consent was not obtained because the data were analyzed retrospectively and anonymously.

Declaration of Competing Interest

All authors declare that they have no financial or other conflicts of interest.

Acknowledgments

The authors gratefully acknowledge An Pan and Zhiguo Zhang for their contributions to the revision of the manuscript and constructive suggestions. We appreciate Dr. Cordelle Lazare and Lai Hsiaw Jie for their contributions in language editing. This work was supported by the Natural Science Foundation of China (81630060 and 81772775), National Basic Research Program (973 Program) from the Ministry of Science and Technology of the People's Republic of China (2015CB553903), and the research-oriented clinician funding program of Tongji Medical College, Huazhong University of Science and Technology. The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygyno.2019.05.026.

References

[1] WHO, Human Papillomavirus (HPV) and Cervical Cancer, 2018.
[2] G.Y. Ho, R.D. Burk, S. Klein, A.S. Kadish, C.J. Chang, P. Palan, et al., Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia, J. Natl. Cancer Inst. 87 (1995) 1365–1371.
[3] S.K. Kjaer, A.J. van den Brule, G. Paull, E.I. Svare, M.E. Sherman, B.L. Thomsen, et al., Type specific persistence of high risk human papillomavirus (HPV) as indicator of high grade cervical squamous intraepithelial lesions in young women: population based prospective follow up study, BMJ 325 (2002) 572.
[4] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (2018) 394–424.

[5] S. Franceschi, L. Denny, K.L. Irwin, J. Jeronimo, P.L. Lopalco, J. Monsonego, et al., Eurogin 2010 roadmap on cervical cancer prevention, Int. J. Cancer 128 (2011) 2765–2774.
[6] W.K. Huh, K.A. Ault, D. Chelmow, D.D. Davey, R.A. Goulart, F.A. Garcia, et al., Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance, Gynecol. Oncol. 136 (2015) 178–182.
[7] R.P. Insinga, G. Perez, C.M. Wheeler, L.A. Koutsky, S.M. Garland, S. Leodolter, et al., Incident cervical HPV infections in young women: transition probabilities for CIN and infection clearance, Cancer Epidemiol. Biomark. Prev. 20 (2011) 287–296.
[8] H. Trottier, S. Mahmud, J.C. Prado, J.S. Sobrinho, M.C. Costa, T.E. Rohan, et al., Type-specific duration of human papillomavirus infection: implications for human papillomavirus screening and vaccination, J. Infect. Dis. 197 (2008) 1436–1447.
[9] M.T. Goodman, Y.B. Shvetsov, K. McDuffie, L.R. Wilkens, X. Zhu, P.J. Thompson, et al., Prevalence, acquisition, and clearance of cervical human papillomavirus infection among women with normal cytology: Hawaii human papillomavirus cohort study, Cancer Res. 68 (2008) 8813–8824.
[10] N. Munoz, G. Hernandez-Suarez, F. Mendez, M. Molano, H. Posso, V. Moreno, et al., Persistence of HPV infection and risk of high-grade cervical intraepithelial neoplasia in a cohort of Colombian women, Br. J. Cancer 100 (2009) 1184–1190.
[11] G.Y. Ho, R. Bierman, L. Beardsley, C.J. Chang, R.D. Burk, Natural history of cervicovaginal papillomavirus infection in young women, N. Engl. J. Med. 338 (1998) 423–428.
[12] C.B. Woodman, S. Collins, H. Winter, A. Bailey, J. Ellis, P. Prior, et al., Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study, Lancet 357 (2001) 1831–1836.
[13] H. Richardson, G. Kelsall, P. Tellier, H. Voyer, M. Abrahamowicz, A. Ferenczy, et al., The natural history of type-specific human papillomavirus infections in female university students, Cancer Epidemiol. Biomark. Prev. 12 (2003) 485–490.
[14] J.E. Koshiol, J.C. Schroeder, D.J. Jamieson, S.W. Marshall, A. Duerr, C.M. Heilig, et al., Time to clearance of human papillomavirus infection by type and human immunodeficiency virus serostatus, Int. J. Cancer 119 (2006) 1623–1629.
[15] K. Louvanto, M.A. Rintala, K.J. Syrjanen, S.E. Grenman, S.M. Syrjanen, Genotype-specific persistence of genital human papillomavirus (HPV) infections in women followed for 6 years in the Finnish family HPV study, J. Infect. Dis. 202 (2010) 436–444.
[16] D.R. Brown, M.L. Shew, B. Qadadri, N. Neptune, M. Vargas, W. Tu, et al., A longitudinal study of genital human papillomavirus infection in a cohort of closely followed adolescent women, J. Infect. Dis. 191 (2005) 182–192.
[17] S.P. Day, A. Hudson, A. Mast, T. Sander, M. Curtis, S. Olson, et al., Analytical performance of the investigational use only Cervista HPV HR test as determined by a multi-center study, J. Clin. Virol. 45 (Suppl. 1) (2009) S63–S72.
[18] K. Louvanto, K.J. Syrjanen, M.A. Rintala, S.E. Grenman, S.M. Syrjanen, Genotype-specific clearance of genital human papillomavirus (HPV) infections among mothers in the Finnish family HPV study, J. Clin. Microbiol. 48 (2010) 2665–2671.
[19] N.W. Bulkmans, J. Berkhof, S. Bulk, M.C. Bleeker, F.J. van Kemenade, L. Rozendaal, et al., High-risk HPV type-specific clearance rates in cervical screening, Br. J. Cancer 96 (2007) 1419–1424.
[20] A.C. Rodriguez, M. Schiffman, R. Herrero, S. Wacholder, A. Hildesheim, P.E. Castle, et al., Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections, J. Natl. Cancer Inst. 100 (2008) 513–517.
[21] M. Molano, A. Van den Brule, M. Plummer, E. Weiderpass, H. Posso, A. Arslan, et al., Determinants of clearance of human papillomavirus infections in Colombian women with normal cytology: a population-based, 5-year follow-up study, Am. J. Epidemiol. 158 (2003) 486–494.
[22] M. Plummer, M. Schiffman, P.E. Castle, D. Maucort-Boulch, C.M. Wheeler, A. Group, A 2-year prospective study of human papillomavirus persistence among women with a cytological diagnosis of atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion, J. Infect. Dis. 195 (2007) 1582–1589.
[23] C. Ingabire, M.K. Lim, Y.J. Won, J.K. Oh, Human papillomavirus genotype-specific persistence and potential risk factors among Korean women: results from a 2-year follow-up study, Cancer Res. Treat. 50 (2018) 813–822.
[24] C.E. Schmeink, W.J. Melchers, A.G. Siebers, W.G. Quint, L.F. Massuger, R.L. Bekkers, Human papillomavirus persistence in young unscreened women, a prospective cohort study, PLoS One 6 (2011), e27937.
[25] A. Hildesheim, M.H. Schiffman, P.E. Gravitt, A.G. Glass, C.E. Greer, T. Zhang, et al., Persistence of type-specific human papillomavirus infection among cytologically normal women, J. Infect. Dis. 169 (1994) 235–240.
[26] S.M. Kulmala, I.P. Shabalova, N. Petrovitchev, K.J. Syrjanen, U.B. Gyllensten, B.C. Johansson, et al., Type-specific persistence of high-risk human papillomavirus infections in the new independent states of the former Soviet Union cohort study, Cancer Epidemiol. Biomark. Prev. 16 (2007) 17–22.
[27] M. Mollers, J. Boot Hein, J. Vriend Henrike, J. King Audrey, V.F. van den Broek Ingrid, E.A. van Bergen Jan, et al., Prevalence, incidence and persistence of genital HPV infections in a large cohort of sexually active young women in the Netherlands, Vaccine 31 (2013) 394–401.
[28] A. Nielsen, T. Iftner, C. Munk, S.K. Kjaer, Acquisition of high-risk human papillomavirus infection in a population-based cohort of Danish women, Sex. Transm. Dis. 36 (2009) 609–615.
[29] E.I. Benchimol, L. Smeeth, A. Guttmann, K. Harron, D. Moher, I. Petersen, et al., The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement, PLoS Med. 12 (2015), e1001885.
[30] E.M. Kurian, M.L. Caporelli, S. Baker, B. Woda, E.F. Cosar, L. Hutchinson, Cervista HR and HPV 16/18 assays vs hybrid capture 2 assay, outcome comparison in women with negative cervical cytology. Am J Clin Pathol. 136 (2011) 808–816.
[31] S.S. Liu, R.C. Leung, K.K. Chan, A.N. Cheung, H.Y. Ngan, Evaluation of a newly developed GenoArray human papillomavirus (HPV) genotyping assay and comparison with the Roche linear Array HPV genotyping assay, J. Clin. Microbiol. 48 (2010) 758–764.
[32] P.D. Allison, Chapter 5: Estimating cox regression models with PROC PHREG, in: G. McDaniel (Ed.), Survival Analysis Using SAS: A Practical Guide, second editionSAS Institute Inc., Cary, NC, USA 2010, pp. 153–179.
[33] David G. Kleinbaum MK. Chapter 6: Extension of the cox proportional hazards model. In: M. Gail KK, J. Samet, A. Tsiatis, W. Wong, editor. Statistics for Biology and Health. Second Edition ed. New York, the U.S.: Springer Science and Business Media Inc.; 2005. p. 224–235.
[34] K. Syrjanen, S.M. Kulmala, I. Shabalova, N. Petrovichev, V. Kozachenko, T. Zakharova, et al., Epidemiological, clinical and viral determinants of the increased prevalence of high-risk human papillomavirus (HPV) infections in elderly women, Eur. J. Gynaecol. Oncol. 29 (2008) 114–122.
[35] C.B. Woodman, S.I. Collins, L.S. Young, The natural history of cervical HPV infection: unresolved issues, Nat. Rev. Cancer 7 (2007) 11–22.
[36] M. Schiffman, P.E. Castle, J. Jeronimo, A.C. Rodriguez, S. Wacholder, Human papillomavirus and cervical cancer, Lancet 370 (2007) 890–907.
[37] K. Sundstrom, A. Ploner, L. Arnheim-Dahlstrom, S. Eloranta, J. Palmgren, H.O. Adami, et al., Interactions between high- and low-risk HPV types reduce the risk of squamous cervical Cancer, J. Natl. Cancer Inst. 107 (2015).
[38] D.C. Beachler, G. Jenkins, M. Safaeian, A.R. Kreimer, N. Wentzensen, Natural acquired immunity against subsequent genital human papillomavirus infection: a systematic review and meta-analysis, J. Infect. Dis. 213 (2016) 1444–1454.
[39] C.E. Banister, A.R. Messersmith, B. Cai, L.B. Spiryda, S.H. Glover, L. Pirisi, et al., Disparity in the persistence of high-risk human papillomavirus genotypes between African American and European American women of college age, J. Infect. Dis. 211 (2015) 100–108.