Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Clinical Setting as a Modifier of HPV Co-infection Epidemiology: Evidence for Fundamentally Different Age–Co-infection Dynamics Between Asymptomatic Screening and Symptomatic Outpatient Populations

Yu Zhang

chrisss1226@163.com

Wuhan Asia Heart Hospital Affiliated to Wuhan University

Erick Thokerunga

Infectious Disease Institute

Christian Cedric Bongolo

Chinese Academy of Science

Research Article

Keywords: human papillomavirus, co-infection, cervical cancer screening, genotype distribution, competitive exclusion

Posted Date: May 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9406286/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Clinical Setting as a Modifier of HPV Co-infection Epidemiology: Evidence for Fundamentally Different Age–Co-infection Dynamics Between Asymptomatic Screening and Symptomatic Outpatient Populations

Yu Zhang[1], Erick Thokerunga[2, 3], Christian Cedric Bongolo[4]

[1] Wuhan Asia Heart Hospital Affiliated to Wuhan University, 430072 Wuhan China

[2] Infectious Disease Institute P.O. Box 22418, Kampala, Uganda

[3] Warr Health Center IV, P.O.BOX 6 Paidha, Zombo, Uganda

[4] CAS Key Lab of Bio-Medical Diagnostics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou, 215163, China

Corresponding authors: Yu Zhang, Wuhan Asia Heart Hospital Affiliated to Wuhan University, 430072 Wuhan China. Eamil: chrisss1226@163.com Dr. Christian Cedric Bongolo, CAS Key Lab of Bio-Medical Diagnostics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou, 215163, China. Cedro91@hotmail.com

Abstract

Background: HPV co-infection with multiple genotypes is epidemiologically and clinically significant, yet whether co-infection patterns differ systematically between asymptomatic screening and symptomatic clinical populations remains poorly characterized.

Methods: We performed a secondary analysis of individual-level cervical HPV genotyping data from 37,967 women in Taizhou, China (2010–2019). Co-infection prevalence, age-stratified patterns, genotype pairing, and hrHPV co-infection composition were compared between asymptomatic and symptomatic cohorts. Binary logistic regression identified independent predictors of co-infection, and a pooled interaction model tested whether the age–co-infection relationship differed by clinical setting.

Results: Co-infection prevalence was substantially higher in the asymptomatic cohort (22.1%) than the symptomatic cohort (7.5%). The asymptomatic cohort showed a monotonic increase in co-infection with age, while the symptomatic cohort exhibited a U-shaped pattern, elevated in both the youngest and oldest women. This divergence was confirmed as a statistically significant effect modification (LRT $\chi^2(5) = 17.62$, $p = 0.0035$). Women aged 55–64 years in the symptomatic cohort had fourfold higher adjusted odds of co-infection (OR 4.29, 95% CI 2.16–8.52). HPV52 dominated both cohorts; HPV16 ranked second in the symptomatic cohort versus fourth in the asymptomatic cohort. Competitive exclusion between dominant hrHPV genotypes was more pronounced in the symptomatic cohort (HPV16+HPV52: O/E = 0.14), and hrHPV+hrHPV dual infections constituted 64.6% of symptomatic versus 43.0% of asymptomatic co-infections.

Conclusion: Clinical setting significantly modifies co-infection patterns and should inform triage protocols and epidemiological surveillance. Peri- and post-menopausal women attending outpatient gynecology services represent a high-priority group for extended genotyping and expedited colposcopy referral.

Keywords: human papillomavirus; co-infection; cervical cancer screening; genotype distribution; competitive exclusion

1. Introduction

Cervical cancer remains a major global public health burden, with an estimated 660,000 new cases and 350,000 deaths recorded worldwide in 2022, the overwhelming majority occurring in low- and middle-income countries [1]. Persistent infection with oncogenic human papillomavirus (HPV) is the necessary cause of virtually all cervical cancers, and the epidemiological transition from HPV infection to invasive malignancy is now well characterized [2,3]. The advent of prophylactic HPV vaccines and the global scale-up of HPV-based cervical screening have generated renewed urgency to understand the distribution and behavior of HPV genotypes at the population level, particularly in high-burden settings where screening infrastructure is rapidly expanding [4,5]. China carries a disproportionate share of the global cervical cancer burden, accounting for approximately 18% of new cases annually, while the epidemiology of HPV infection in Chinese women differs meaningfully from Western populations, with HPV52 and HPV58 assuming prominence alongside HPV16 and HPV18 as dominant oncogenic genotypes [6,7]. Accurate characterization of genotype-specific prevalence in Chinese populations is therefore essential both for informing vaccine policy decisions and for optimizing the clinical performance of genotype-based triage algorithms within national screening programmes [8].

Within the broader landscape of HPV epidemiology, co-infection with multiple HPV genotypes represents a clinically and virologically important but incompletely understood phenomenon. Pooled analyses of large international prevalence surveys have demonstrated that concurrent multi-genotype infection is common, occurring in 10–30% of HPV-positive women depending on the population studied and the genotyping platform used [9,10]. Co-infection is not epidemiologically neutral: several prospective studies have reported that women harboring dual high-risk HPV genotypes face a higher risk of high-grade cervical intraepithelial neoplasia and invasive cancer than women with single high-risk infections, suggesting a potential cumulative oncogenic effect [11,12]. Furthermore, the pairwise co-occurrence of specific HPV genotypes deviates systematically from random expectation, with competitive exclusion (defined as significantly lower-than-expected co-occurrence) documented between the most prevalent high-risk types in multiple independent datasets [13,14]. The biological mechanisms underlying this competitive exclusion remain contested, with proposed explanations including type-specific immune cross-reactivity, competition for susceptible epithelial cell receptor sites, and interference at the level of viral replication and episomal maintenance [15,16]. Resolving these questions has practical implications for understanding how vaccination-induced reductions in HPV16 and HPV18 prevalence may alter the ecological niche available to other oncogenic genotypes — the so-called type-replacement hypothesis — which remains an active area of post-licensure vaccine surveillance [17].

A critically underexplored dimension of HPV co-infection epidemiology is whether co-infection patterns differ systematically between women attending routine population-based screening and those presenting to clinical services with gynecological symptoms or on referral. These two populations occupy fundamentally different positions on the cervical disease continuum: asymptomatic screening populations are enriched for prevalent but immunologically controlled infections, while symptomatic clinical populations are more likely to harbor infections associated with progressive cervical pathology [18,19]. If

co-infection prevalence, genotype composition, and age patterning differ between these settings, then estimates derived from one population may not be directly applicable to the other, with implications for clinical triage algorithms, referral thresholds, and the interpretation of cohort-specific epidemiological data [20]. Yet comparative analyses explicitly designed to quantify these between-setting differences are scarce in the literature, and those that exist have generally been limited by small sample sizes, incomplete genotyping panels, or the absence of formal statistical testing for setting-by-age interaction effects [21,22].

The present study addresses this gap by performing a comparative secondary analysis of individual-level genotyping data from 37,967 women in Taizhou, Zhejiang Province, China. We compare HPV co-infection prevalence, genotype pairing patterns, age-stratified co-infection dynamics, and predictors of multi-genotype infection between a Physical Examination (PE) cohort of asymptomatic women attending routine cervical cancer screening (n = 10,068) and an Outpatient Department (OPD) cohort of women attending gynecology clinics (n = 27,899). A pooled logistic regression framework with formal interaction testing is used to determine whether the age–co-infection relationship differs significantly between clinical settings. The findings have direct relevance to the design of HPV triage protocols, the targeting of extended genotyping panels, and the interpretation of HPV epidemiological surveillance data across diverse clinical contexts in China and comparable high-burden settings.

2. Methods

2.1. Study Design and Data Source

This is a secondary analysis of anonymized individual-level data from a previously published cross-sectional study examining the prevalence and characteristics of cervical HPV genotypes among women in the Taizhou area, Zhejiang Province, China [23]. The primary dataset, comprising 37,967 women recruited between 2010 and 2019, was obtained from the Dryad Digital Repository (https://doi.org/10.5061/dryad.c7h3c). Ethics approval and written informed consent for the original data collection were obtained by the primary investigators from the Institutional Medical Ethics Review Board of Taizhou Hospital in Zhejiang Province. Our secondary use of the anonymized dataset did not require additional ethics review. This study was conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki.

The present analysis encompasses two sub-cohorts drawn from this dataset. The Physical Examination (PE) cohort (n = 10,068) comprises women who attended routine asymptomatic cervical cancer screening programmes. The Outpatient Department (OPD) cohort (n = 27,899) comprises women who attended outpatient gynecology or related clinics, many presenting with gynecological symptoms or on referral. Comparative analysis of these two clinically distinct populations enables characterization of HPV co-infection burden across the screening–symptomatic care continuum.

2.2. Study Population and Eligibility

PE cohort: All women aged 21–82 years attending physical examination clinics who underwent cervical HPV genotype testing. **OPD cohort**: all women aged 15–90 years attending outpatient clinics who

underwent testing. Exclusion criteria applied in the original study included prior hysterectomy, current pregnancy, inability to provide informed consent, and active menstruation at the time of sampling.

2.3. Cervical HPV Genotyping

Cervical specimens were collected by trained gynecologists using a standardized Dacron swab. HPV genotyping was performed using the flow-through hybridization and gene chip method (HybriMax®; Yaneng Biosciences, Shenzhen, China), capable of simultaneously detecting 21 HPV genotypes. Genotypes were further classified per International Agency for Research on Cancer (IARC) Working Group criteria: Group 1 carcinogens (definite high-risk; hrHPV): HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68; Group 2A/2B (probable/possible high-risk; pHR): HPV 26, 53, 73, and 82; and low-risk types (lrHPV): all remaining detectable genotypes.

2.4. Variable Definitions

HPV co-infection was defined as detection of two or more distinct HPV genotypes in a single cervical sample. Infection multiplicity was sub-classified as single infection (one genotype), double (two genotypes), triple (three genotypes), and quadruple or higher (≥4 genotypes). Age groups were defined a priori as: <25, 25–34, 35–44, 45–54, and ≥55 years, consistent with HPV epidemiological literature in China. For regression analyses, the ≥55 group was further split into 55–64 and ≥65 years to capture finer age-related trends.

2.5. Statistical Analysis

Prevalence was expressed as proportions with 95% confidence intervals (Clopper–Pearson exact method). Age-group differences in co-infection rates were assessed by Pearson chi-square and the Cochran–Armitage test for linear trend. Spearman's rank correlation quantified the monotonic age–co-infection association. Pairwise genotype co-occurrence was assessed against expected frequencies under the independence assumption [$E(A,B) = P(A) \times P(B) \times N$] using chi-square goodness-of-fit tests. $O/E > 1.2$ was considered positive affinity; $O/E < 0.8$ co-occurrence avoidance. Binary logistic regression identified independent predictors of co-infection among HPV-positive women; covariates were age (continuous, per year) and age group (categorical; reference: 35–44 years). Unadjusted and adjusted odds ratios (OR) with 95% CIs are reported. Model fit was assessed using the Hosmer–Lemeshow test and Nagelkerke pseudo-R^2.

A pooled binary logistic regression was conducted to assess whether the relationship between age and HPV co-infection differed between PE and OPD cohorts. The analysis included 8,549 HPV-positive women aged 21–82 years, combining PE (n = 1,963) and OPD (n = 6,586) datasets with a cohort indicator. Three hierarchical models were fitted: Model A included age group only, Model B added cohort, and Model C incorporated age group–cohort interaction terms to test effect modification. Age-specific odds ratios within each cohort were estimated from Model C using the delta method. Model performance and interaction effects were evaluated using likelihood ratio tests comparing Models B and C, with significance assessed via a chi-squared distribution (df = 5). Model calibration was examined using the Hosmer–Lemeshow test. A sensitivity analysis including the full OPD age range (15–90 years) confirmed

robustness. All analyses were performed in Python 3.11 (Pandas, NumPy, SciPy, statsmodels). All tests were two-sided ($\alpha = 0.05$).

3. Results

3.1. Study Population

Physical Examination cohort: A total of 10,068 women were included, aged 21–82 years (median 42 years). The majority (70.6%) were aged 35–54 years. Of the total, 1,964 women (19.5%) tested HPV-positive.

OPD cohort: A total of 27,899 women were included, aged 15–90 years (median 41 years). The OPD cohort was younger on average, with a higher proportion of women aged <35 years (28.7% vs. 16.3% in the PE cohort). Of the total, 6,705 women (24.0%) tested HPV-positive, a significantly higher positivity rate than in the PE cohort.

The full demographic and infection characteristics of both cohorts, including age distribution and HPV positivity by stratum, are presented in **Table 1**. The age-specific HPV positivity curves for both cohorts, smoothed with a 5-point moving average, are shown in **Figure 1**. The PE cohort maintains a relatively stable positivity rate across the adult lifespan, with only modest variation between age groups. In contrast, the OPD cohort shows markedly elevated positivity in women under 25 years (exceeding 40% in the youngest age years) and a pronounced secondary rise around ages 55–65, reflecting the distinct clinical and demographic profile of women attending outpatient gynecology services compared with those presenting for routine asymptomatic screening.

Table 1. Demographic and HPV infection characteristics of the study population, by cohort.

Characteristic	Physical Examination Cohort (n = 10,068)			Outpatient Department Cohort (n = 27,899)		
	n (%)	HPV-positive n (%)	Co-infection rate (%)	n (%)	HPV-positive n (%)	Co-infection rate (%)
Overall	10,068 (100)	1,964 (19.5)	434 (22.1)	27,899 (100)	6,705 (24.0)	505 (7.5)
Age group (years)						
<25	36 (0.4)	9 (25.0)	5 (55.6)	1,187 (4.3)	361 (30.4)	36 (10.0)
25–34	1,597 (15.9)	292 (18.3)	60 (20.5)	6,810 (24.4)	1,507 (22.1)	119 (7.9)
35–44	3,595 (35.7)	653 (18.2)	124 (19.0)	10,197 (36.6)	2,298 (22.5)	137 (6.0)
45–54	3,513 (34.9)	729 (20.8)	161 (22.1)	7,470 (26.8)	1,841 (24.6)	123 (6.7)
≥55	1,327 (13.2)	281 (21.2)	84 (29.9)	2,235 (8.0)	698 (31.2)	90 (12.9)
Infection multiplicity (among HPV-positive)						
Single (n = 1 genotype)	–	1,530 (77.9)	–	–	6,200 (92.5)	–
Double (n = 2 genotypes)	–	336 (17.1)	–	–	385 (5.7)	–
Triple (n = 3 genotypes)	–	67 (3.4)	–	–	85 (1.3)	–
Quadruple or higher (≥4 genotypes)	–	31 (1.6)	–	–	35 (0.5)	–

Abbreviations: HPV+ = HPV-positive; co-infection = detection of ≥2 HPV genotypes. Co-infection rate is expressed as a percentage of HPV-positive women within each stratum. Dash (–) indicates not applicable.

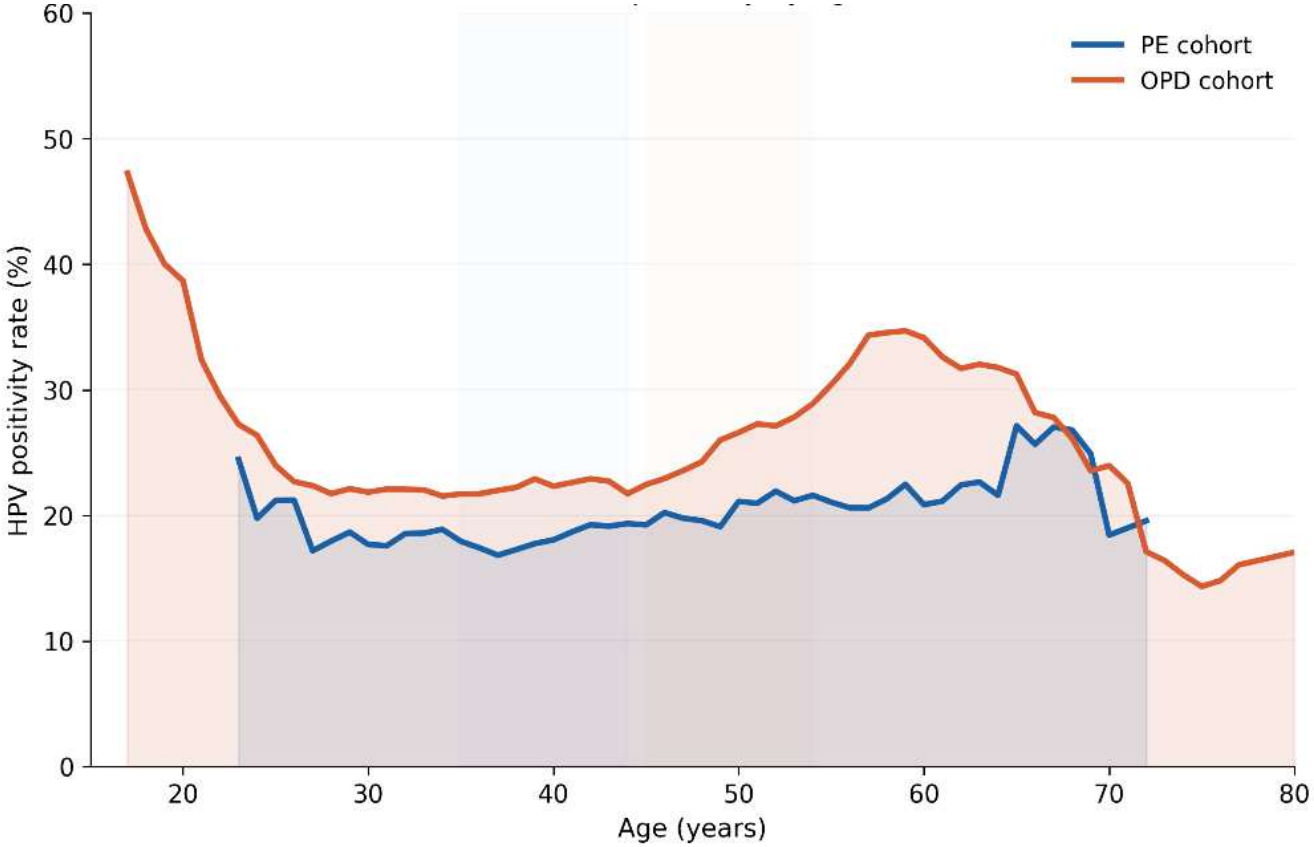


Figure 1. Age-prevalence curves by cohort (5-point moving average). HPV positivity rate by single year of age in the Physical Examination (PE) and Outpatient Department (OPD) cohorts. Shaded regions denote the age range common to both cohorts (21–82 years). The PE cohort shows a monotonically increasing co-infection trend with age; the OPD cohort shows a U-shaped pattern with nadirs at 35–44 years.

3.2. Overall Co-infection Prevalence

Physical Examination cohort: Among 1,964 HPV-positive women, 1,530 (77.9%) had single-genotype infections and 434 (22.1%) had co-infections with ≥2 genotypes. Double co-infections were most common (n = 336; 17.1%), followed by triple (n = 67; 3.4%) and quadruple-or-higher (n = 31; 1.6%). Approximately one in five HPV-positive women in this asymptomatic screening population harbored concurrent multi-genotype infections.

OPD cohort: Among 6,705 HPV-positive women, 6,200 (92.5%) had single-genotype infections and 505 (7.5%) had co-infections — a co-infection rate nearly three times lower than in the PE cohort (7.5% vs. 22.1%). Of co-infected women, the distribution was: double (n = 385; 5.7%), triple (n = 85; 1.3%), and quadruple-or-higher (n = 35; 0.5%). The substantially lower absolute co-infection rate in the OPD cohort, despite higher overall HPV positivity, likely reflects the much larger denominator of HPV-positive women in this cohort and the difference in the genotyping pattern between symptomatic and asymptomatic women.

3.3. Age-Stratified Co-infection

Physical Examination cohort: Co-infection rates varied significantly across age groups ($\chi^2 = 19.85$, df = 4, $p < 0.001$). Among larger strata, rates were lowest in the 35–44 year group (19.0%) and rose with advancing age to 29.9% in women aged ≥55 years. The Cochran–Armitage test confirmed a significant linear trend with age ($p = 0.002$), and Spearman's correlation indicated a positive association ($\rho = 0.070$, $p = 0.002$).

OPD cohort: Co-infection rates also varied significantly across age groups ($\chi^2 = 42.24$, df = 4, $p < 0.001$), but followed a U-shaped pattern rather than a linear trend. Rates were highest in the youngest women (<25 years: 10.0%) and in the oldest (age ≥55 years: 12.9%), with the nadir at 35–44 years (6.0%). Unlike the PE cohort, the Cochran–Armitage test did not confirm a significant linear trend ($p = 0.226$), and Spearman's correlation was non-significant ($\rho = 0.012$, $p = 0.331$), indicating that the elevated co-infection risk in OPD attendees is concentrated at both ends of the age spectrum rather than increasing monotonically.

The divergence between these two age patterns is illustrated in **Figure 2,** which plots the 5-point moving average co-infection rate against individual age years for each cohort. The PE cohort (solid blue line) shows a broadly rising trajectory from the mid-30s onward, consistent with the monotonic trend confirmed statistically. The OPD cohort (dashed orange line) traces a visually distinct U-shape, with an early peak in women under 25, a trough centered on the 35–44 year nadir, and a secondary rise among women aged 55 years and older. This graphical representation underscores that the two cohorts do not merely differ in overall co-infection burden but follow fundamentally different age-dependent infection dynamics — a divergence that is formally tested in the pooled interaction analysis presented in Section 3.8.

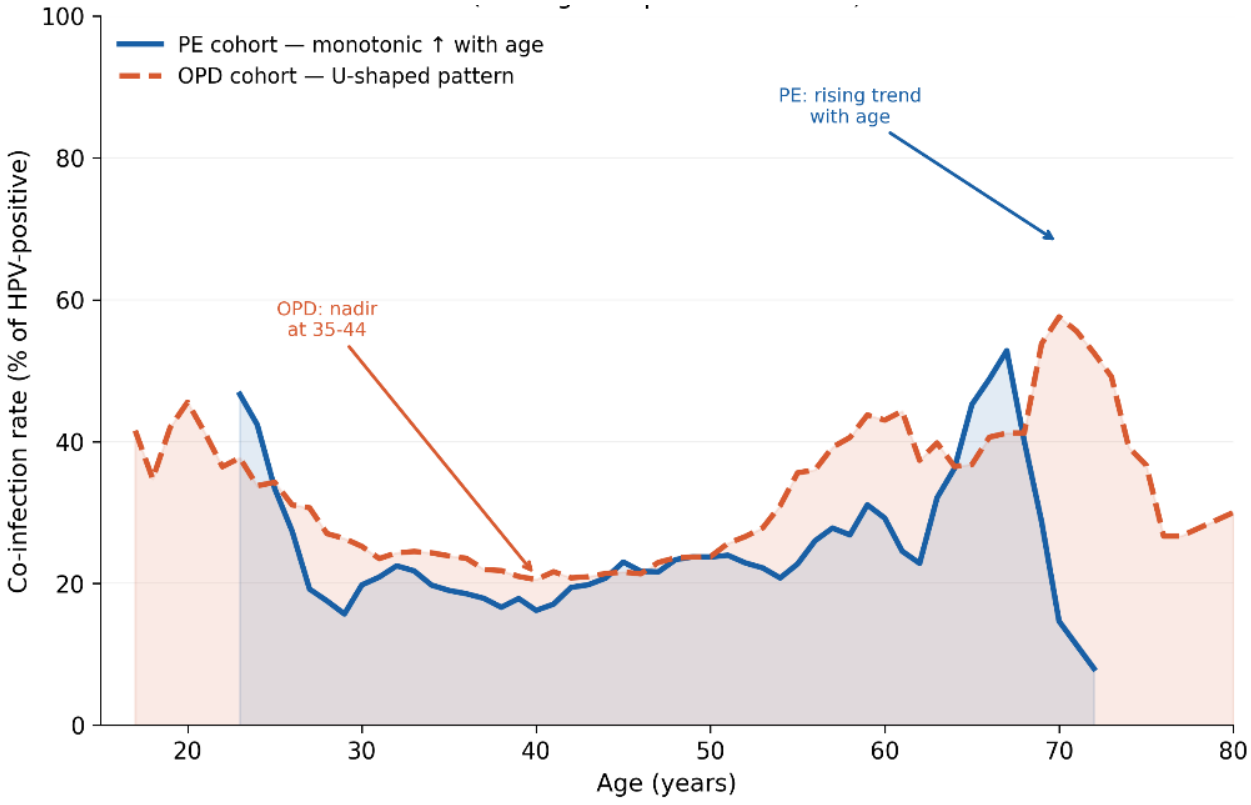


Figure 2. Age-prevalence curves by cohort (5-point moving average). Co-infection rate (proportion of HPV-positive women with ≥2 genotypes) by single year of age in each cohort. All curves are smoothed using a 5-point moving average. Shaded regions denote the age range common to both cohorts (21–82 years). The PE cohort shows a monotonically increasing co-infection trend with age; the OPD cohort shows a U-shaped pattern with nadirs at 35–44 years.

3.4. Genotype Frequency

Physical Examination cohort: HPV52 was the most prevalent genotype (17.4%), followed by HPV58 (10.4%) and HPV53 (10.4%). Among definite high-risk types, the rank order was HPV52 > HPV58 > HPV16 > HPV39 > HPV18. HPV16 ranked fourth (8.2%), below its typical rank in Western populations. HPV52 and HPV58 together accounted for 27.8% of all HPV-positive cases, consistent with East Asian HPV epidemiology.

OPD cohort: HPV52 again dominated (18.5%). Notably, HPV16 ranked second (12.6%) in the OPD cohort versus fourth in the PE cohort, suggesting that clinical-encounter populations are enriched for the highest-oncogenic-risk genotypes. HPV58 ranked third (9.5%), followed by HPV39 (6.1%) and HPV18 (6.0%). The higher HPV16 prevalence in the OPD cohort is clinically significant given its strongest causal association with cervical cancer.

The divergent genotype rank orders between the two cohorts are visually apparent in Figure 3, which displays prevalence as a back-to-back horizontal bar chart with PE cohort bars extending leftward and OPD cohort bars rightward. The dominance of HPV52 in both cohorts is immediately evident from the longest bars at the bottom of the chart, while the disproportionate rightward extension of the HPV16 bar — relative to its PE counterpart — captures the rank elevation of this genotype in the symptomatic outpatient population. Genotypes are colour-coded by IARC risk class, allowing rapid visual identification of the high-risk genotype cluster that accounts for the majority of detections in both cohorts.

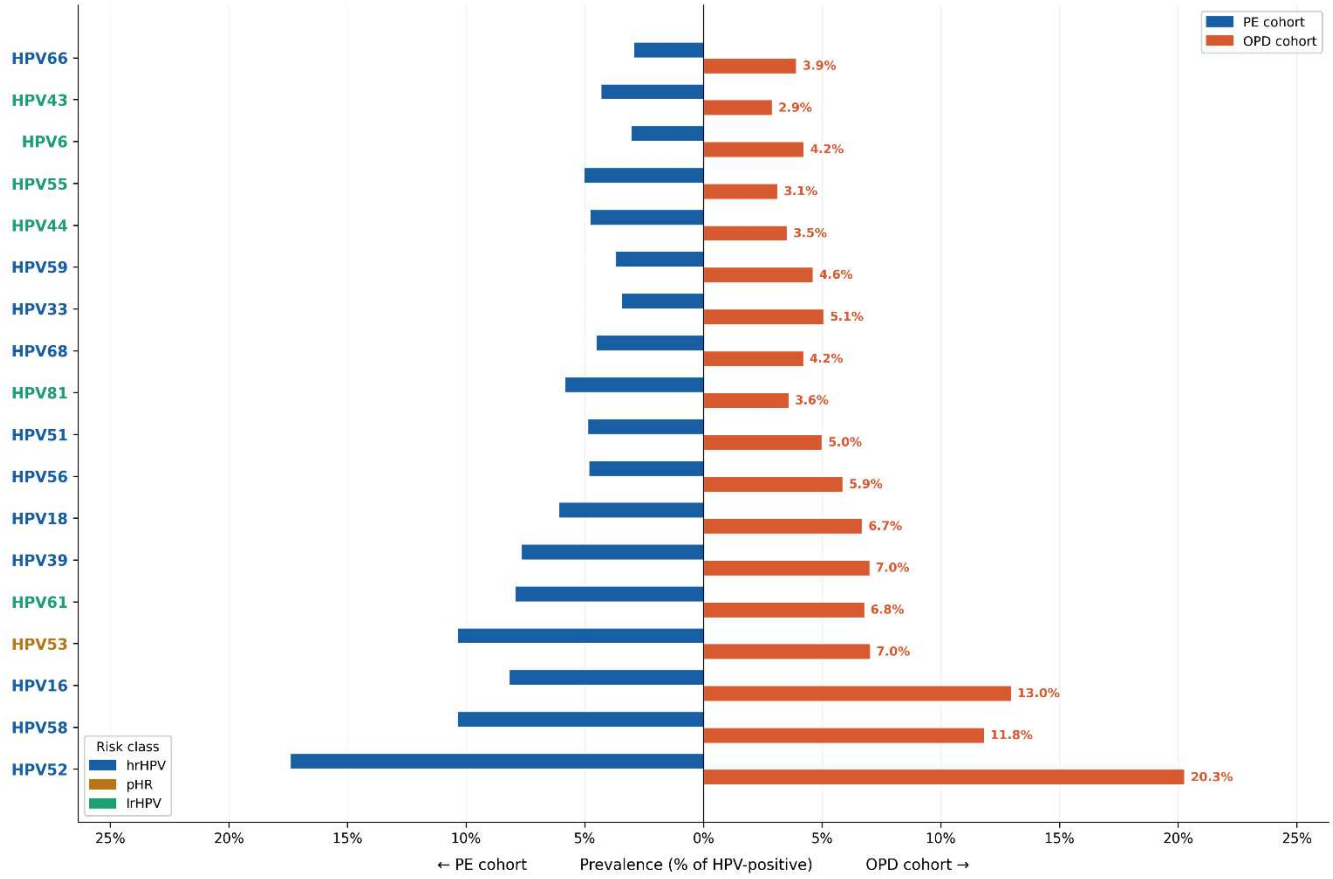


Figure 3. HPV genotype prevalence comparison between the Physical Examination (PE) and Outpatient Department (OPD) cohorts. Bars represent the prevalence of each detected genotype as a percentage of HPV-positive women, displayed as a back-to-back horizontal chart: PE cohort extending leftward (blue) and OPD cohort rightward (orange). Genotypes are ordered by combined prevalence across both cohorts. Colour coding denotes IARC risk classification: high-risk (hrHPV, blue/dark), probable high-risk (pHR, gold), and low-risk (lrHPV, green). OPD percentage labels are shown for selected genotypes. Note the elevation of HPV16 to second rank in the OPD cohort compared with fourth rank in the PE cohort.

3.5. Genotype Pairing Analysis

Physical Examination cohort: The most frequent co-infecting pairs were HPV52+HPV58 (n = 17), HPV52+HPV53 (n = 16), and HPV16+HPV52 (n = 13). Despite their high observed counts, several top pairs occurred significantly less than expected under the independence assumption: HPV52+HPV58 (O/E = 0.48, p = 0.002), HPV52+HPV53 (O/E = 0.45, p = 0.001), and HPV16+HPV52 (O/E = 0.46, p = 0.005), indicating competitive exclusion between the most prevalent genotypes. Conversely, HPV39+HPV43 (O/E = 1.38) showed above-expected co-occurrence.

OPD cohort: The same HPV52+HPV58 pair was the most frequent (n = 30), but competitive exclusion was even stronger (O/E = 0.26, $p < 0.001$). All top ten observed pairs occurred substantially less than expected, with the most extreme being HPV16+HPV52 (O/E = 0.14, $p < 0.001$), suggesting that the two

dominant genotypes in this cohort — HPV52 and HPV16 — show the strongest competitive exclusion. This pattern of systematic co-occurrence avoidance among the most prevalent hrHPV types was more pronounced in the OPD cohort, possibly reflecting greater competitive viral interference in a more genotypically diverse population. The top ten co-infecting pairs with observed and expected frequencies, O/E ratios, and p-values are presented in Table 2.

The full landscape of genotype co-occurrence across all 15 × 15 pairwise combinations is displayed in Figure 4, which plots O/E ratios rather than raw counts, thereby correcting for the confounding effect of individual genotype prevalence on observed pair frequencies. The diverging color scale — red for co-occurrence avoidance (O/E < 0.80), white for random pairing (O/E ≈ 1.0), and blue for positive affinity (O/E > 1.20) — immediately reveals that competitive exclusion, not affinity, is the dominant signal in both cohorts. In Panel A, the PE cohort matrix shows a broadly red field among the high-prevalence hrHPV genotypes in the upper-left quadrant, punctuated by isolated blue cells indicating affinity among lower-prevalence type pairs. Panel B shows an even more uniformly red upper-left quadrant in the OPD cohort, with the HPV52–HPV16 cell among the deepest red in the entire matrix (O/E = 0.14), visually reinforcing the statistical finding that competitive exclusion intensifies in the symptomatic outpatient population. The asterisk annotations within each cell provide direct significance grading, allowing rapid identification of pairs where co-occurrence avoidance or affinity is statistically confirmed rather than merely a product of sampling variation.

Table 2. Top ten co-infecting genotype pairs: observed and expected frequencies, O/E ratio, and p-values — by cohort.

Physical Examination Cohort					Outpatient Department Cohort				
Genotype pair	**Obs**	**Exp**	**O/E**	**p-value**	**Genotype pair**	**Obs**	**Exp**	**O/E**	**p-value**
HPV52+HPV58	17	35.5	0.48	0.002	HPV52+HPV58	30	117.5	0.26	<0.001
HPV52+HPV53	16	35.5	0.45	0.001	HPV52+HPV68	23	40.6	0.57	0.006
HPV16+HPV52	13	28.0	0.46	0.005	HPV16+HPV52	22	156.1	0.14	<0.001
HPV52+HPV68	13	15.5	0.84	0.526	HPV52+HPV56	21	58.4	0.36	<0.001
HPV55+HPV58	12	10.3	1.17	0.592	HPV39+HPV52	18	75.6	0.24	<0.001
HPV44+HPV52	11	16.4	0.67	0.185	HPV16+HPV18	17	50.5	0.34	<0.001
HPV39+HPV81	10	8.8	1.13	0.697	HPV52+HPV61	17	54.5	0.31	<0.001
HPV39+HPV53	10	15.7	0.64	0.151	HPV33+HPV52	15	56.2	0.27	<0.001
HPV53+HPV55	10	10.3	0.97	0.930	HPV52+HPV59	15	43.4	0.35	<0.001
HPV53+HPV58	10	21.2	0.47	0.015	HPV58+HPV59	14	22.1	0.63	0.084

Expected counts: $E(A,B) = P(A) \times P(B) \times N$ under the null hypothesis of independent infection. p-values from chi-square goodness-of-fit test (1 df). Obs = observed count; Exp = expected count; O/E = observed-to-expected ratio.

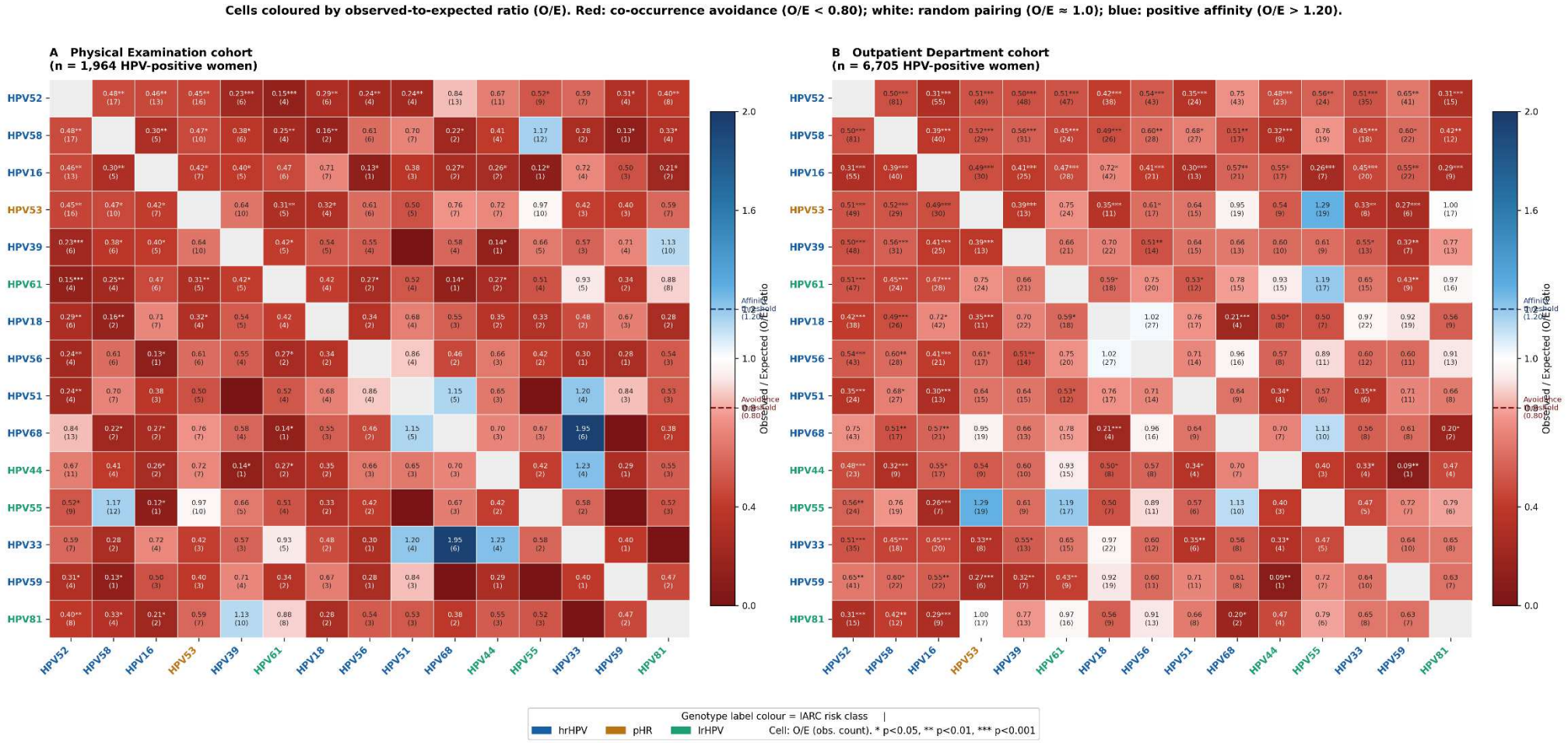


Figure 4. Genotype co-occurrence O/E ratio matrices — competitive exclusion vs. affinity. (A) Physical Examination cohort (n = 1,964 HPV-positive women); (B) Outpatient Department cohort (n = 6,705 HPV-positive women). Each cell displays the observed-to-expected (O/E) ratio with the observed co-occurrence count in parentheses, for all 15 × 15 genotype pairs. Cells are coloured by a diverging scale: red indicates co-occurrence avoidance (O/E < 0.80); white indicates random pairing (O/E ≈ 1.0); blue indicates positive affinity (O/E > 1.20). Dashed horizontal lines mark the O/E = 0.80 and O/E = 1.20 thresholds. Genotype axis labels are colour-coded by IARC risk class: hrHPV (blue), pHR (gold), lrHPV (green). Asterisks denote statistical significance of deviation from independence: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (chi-square goodness-of-fit test). Expected counts calculated as $E(A,B) = P(A) \times P(B) \times N$ under the null hypothesis of independent co-infection.

3.6. hrHPV Co-infection

Physical Examination cohort: Of 1,964 HPV-positive women, 1,344 (68.4%) carried at least one definite hrHPV genotype. Dual or multiple hrHPV co-infections were detected in 185 women (9.4%), and 184 (9.4%) had mixed hrHPV/lrHPV co-infections. Among co-infected women, 43.0% had hrHPV+hrHPV, 41.7% hrHPV+lrHPV, and 14.3% lrHPV+lrHPV co-infections. Dual hrHPV co-infection peaked at 11.4% among women aged ≥55 years.

OPD cohort: Of 6,705 HPV-positive women, 5,050 (75.3%) carried at least one hrHPV genotype — a higher proportion than the PE cohort (75.3% vs. 68.4%), consistent with the symptomatic/referred clinical profile. Dual or multiple hrHPV co-infections were detected in 326 women (4.9%). Among co-infected women in the OPD cohort, 64.6% had hrHPV+hrHPV co-infections, 44.4% mixed hrHPV+lrHPV/pHR infections, and only 2.8% lrHPV+lrHPV co-infections.

These between-cohort differences in co-infection composition are depicted in Figure 5. Panel A shows the prevalence of individual hrHPV genotypes in each cohort, confirming the dominance of HPV52 and the relatively higher burden of HPV16 in the OPD cohort, consistent with the genotype frequency findings in Section 3.4. Panel B illustrates the striking shift in co-infection pair risk class between cohorts: the hrHPV+hrHPV category accounts for 53.8% of co-infected OPD women compared with 42.6% in the PE cohort, while lrHPV+lrHPV co-infections are correspondingly less common in the OPD setting (16.0% vs. 22.6%). Panel C displays dual hrHPV co-infection rates by age group, revealing that OPD women consistently carry a higher dual hrHPV burden across all age strata, with the most pronounced absolute excess in the youngest (<25 years: 22.4% OPD vs. 11.1% PE) and oldest (≥55 years: 22.5% OPD vs. 14.6% PE) women. The substantially higher proportion of hrHPV+hrHPV co-infections in OPD versus PE co-infected women represents a clinically important finding, as dual high-risk genotype infection may confer greater cumulative oncogenic risk.

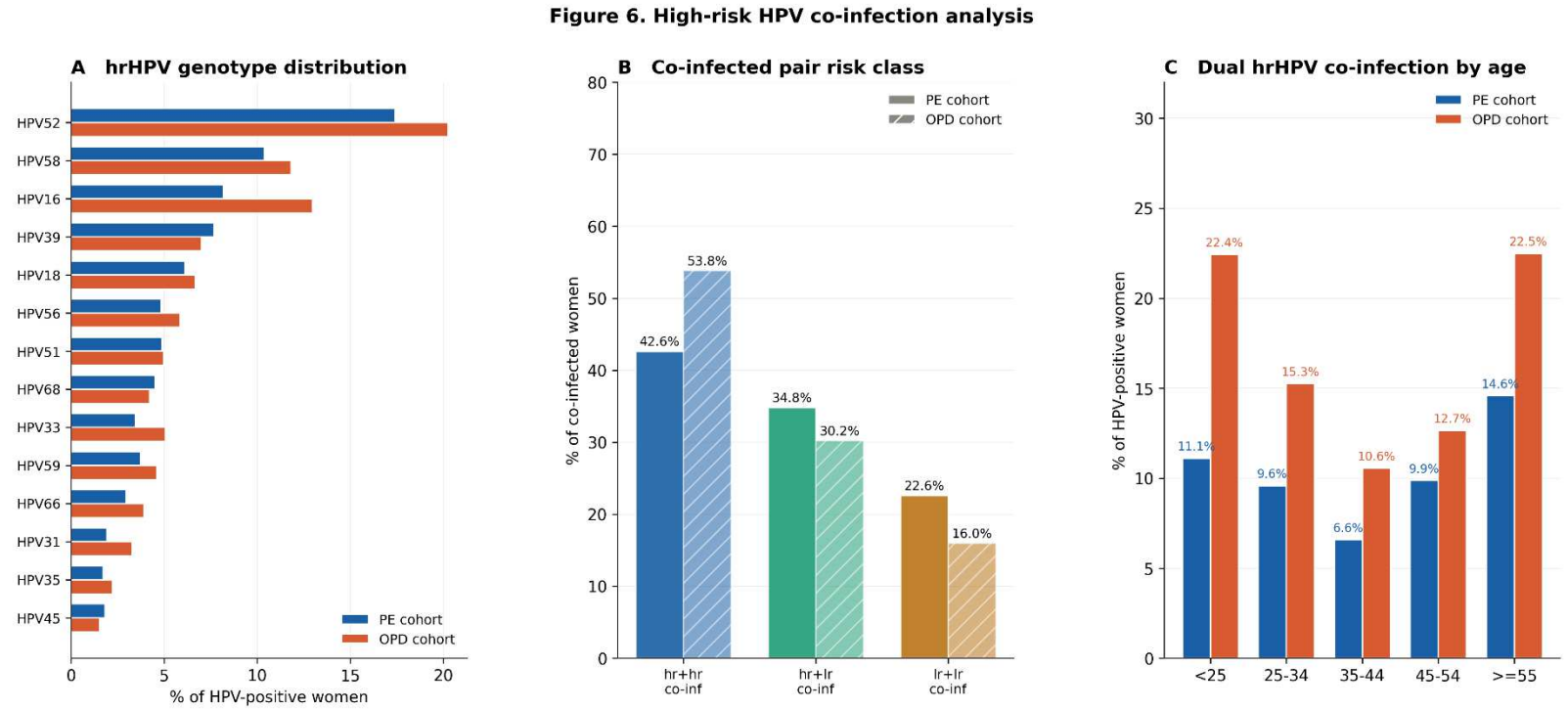


Figure 5. High-risk HPV co-infection analysis by cohort. (A) Prevalence of individual hrHPV genotypes as a percentage of HPV-positive women in the Physical Examination (PE, blue) and Outpatient Department (OPD, orange) cohorts, ordered by PE cohort prevalence. (B) Distribution of co-infection pair risk class among co-infected women: hrHPV+hrHPV, hrHPV+lrHPV (including pHR), and lrHPV+lrHPV categories shown for each cohort. Solid bars = PE cohort; hatched bars = OPD cohort. (C) Dual hrHPV co-infection rate by age group, expressed as a percentage of all HPV-positive women within each stratum. Percentage labels are shown above each bar.

3.7. Predictors of HPV Co-infection: Binary Logistic Regression

Physical Examination cohort: Of the 1,964 HPV-positive women, 31 (1.58%) had co-infections. Nine women aged ≤24 years (zero co-infection events) were excluded due to complete separation, leaving 1,955 in the analytic sample. Age as a continuous variable was a significant unadjusted predictor (OR 1.05, 95% CI 1.01–1.09, p = 0.008). Categorically, only the ≥65 group was significantly associated with co-infection (unadjusted OR 5.11, 95% CI 1.33–19.59, p = 0.017), with a visible increasing trend across strata. In the adjusted model, both predictors lost significance due to collinearity between the two age representations. The overall models showed adequate fit (Hosmer–Lemeshow p > 0.05 for all; Nagelkerke R^2 = 0.024–0.033).

OPD cohort: Of the 6,705 HPV-positive women, 505 (7.5%) had co-infections. All age strata had events, so no group was excluded. Age as a continuous variable was not significant in either the unadjusted (OR 1.00, p = 0.307) or adjusted (OR 0.97, p = 0.098) model, and Spearman's correlation confirmed no linear age–co-infection relationship. However, the categorical age group analysis revealed distinct patterns: the 55–64 year group had the strongest and most clinically meaningful association, with an adjusted OR of

4.29 (95% CI 2.16–8.52, $p < 0.001$) — a finding that persisted and strengthened after controlling for continuous age. The <25 and 25–34 year groups were significantly elevated in unadjusted models but attenuated after adjustment, indicating confounding. The adjusted model met goodness-of-fit criteria (Hosmer–Lemeshow $p = 0.056$; Nagelkerke $R^2 = 0.016$). The full regression output for both cohorts, including unadjusted and adjusted odds ratios for all age strata, is presented in Table 3.

The adjusted odds ratios for both cohorts are displayed as forest plots in Figure 6. Panel A illustrates the wide confidence intervals characteristic of the PE cohort adjusted model, reflecting the small number of co-infection events (n = 31) and the resulting imprecision of age-group estimates — the ≥65 point estimate of OR 8.10 carries a confidence interval spanning nearly three orders of magnitude (0.19–350.90), underscoring the need for cautious interpretation. In contrast, Panel B shows the substantially more precise OPD estimates, with the 55–64 year group standing out as the only age stratum with a confidence interval entirely to the right of the null (OR 4.29, 95% CI 2.16–8.52), visually confirming it as the single most robust and clinically actionable finding of the cohort-specific regression analyses.

Table 3. Binary logistic regression analysis of predictors of HPV co-infection among HPV-positive women, by cohort.

Covariate	Physical Examination Cohort					Outpatient Department Cohort				
	n (ev/tot)	Unadjusted OR (95% CI)	p	Adjusted OR (95% CI)	p	n (ev/tot)	Unadjusted OR (95% CI)	p	Adjusted OR (95% CI)	p
Age (per year increment)										
	31/1,955	1.05 (1.01–1.09)	0.008	0.98 (0.87–1.12)	0.798	505/6,705	1.00 (1.00–1.01)	0.307	0.97 (0.94–1.01)	0.098
Age group (years) — reference: 35–44										
35–44 (Reference)	9/653	1.00 (ref)	–	1.00 (ref)	–	137/2,298	1.00 (ref)	–	1.00 (ref)	–
<25[1]	excl.[2]	—	—	—	—	36/361	1.75 (1.19–2.57)	0.005	1.03 (0.50–2.14)	0.936
25–34	1/292	0.25 (0.03–1.95)	0.184	0.21 (0.02–2.26)	0.199	119/1,507	1.35 (1.05–1.74)	0.020	1.01 (0.66–1.55)	0.970
45–54	12/729	1.20 (0.50–2.86)	0.685	1.40 (0.32–6.01)	0.654	123/1,841	1.13 (0.88–1.45)	0.343	1.45 (0.98–2.14)	0.061
55–64	6/236	1.87 (0.66–5.30)	0.241	2.55 (0.19–34.01)	0.480	82/592	2.54 (1.90–3.39)	<0.001	4.29 (2.16–8.52)	<0.001
≥65	3/45	5.11 (1.33–19.59)	0.017	8.10 (0.19–350.9)	0.277	8/106	1.29 (0.61–2.70)	0.504	2.93 (0.86–9.91)	0.085

*[1] <25 age group in PE cohort excluded due to complete separation (n = 9, zero co-infection events). [2] — indicates covariate not estimable. Unadjusted models entered each covariate individually; adjusted model included both age (continuous) and age group simultaneously. The ≥65 group in the PE cohort (n = 45, events = 3) should be interpreted with caution due to small event counts. * p < 0.05; ** p < 0.001.*

Model fit – PE cohort: Unadj. age continuous: HL $\chi^2(8) = 7.644$, p = 0.469, Nagelkerke $R^2 = 0.024$; Unadj. age group: HL $\chi^2(2) = 0.000$, p = 1.000, $R^2 = 0.033$; Adjusted: HL $\chi^2(8) = 12.550$, p = 0.128, $R^2 = 0.033$.

Model fit – OPD cohort: Unadj. age continuous: HL $\chi^2(8) = 51.775$, p < 0.001, Nagelkerke $R^2 = 0.000$; Unadj. age group: HL $\chi^2(2) = 0.000$, p = 1.000, $R^2 = 0.015$; Adjusted: HL $\chi^2(8) = 15.174$, p = 0.056, $R^2 = 0.016$.

Abbreviations: OR, odds ratio; CI, confidence interval; ref, reference category; ev/tot, events/total; excl., excluded.

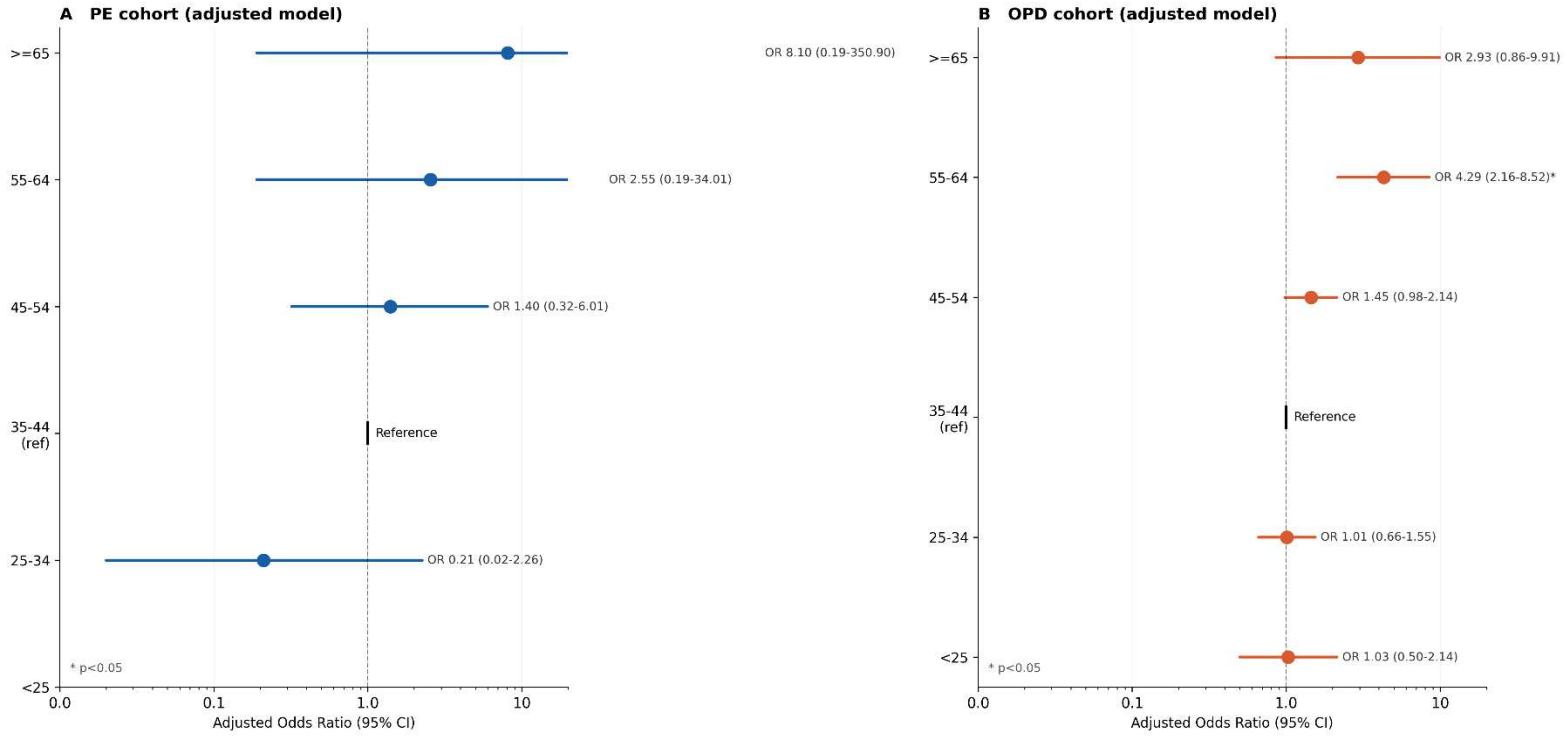


Figure 6. Forest plots of adjusted odds ratios for HPV co-infection by age group, from binary logistic regression models. (A) Physical Examination (PE) cohort; (B) Outpatient Department (OPD) cohort. Points represent adjusted odds ratios; horizontal lines represent 95% confidence intervals. The reference category is women aged 35–44 years (shown as a vertical tick at OR = 1.0). The x-axis is on a log scale. The <25 age group is excluded from Panel A due to complete separation (zero co-infection events). Asterisk (*) denotes $p < 0.05$. OR = odds ratio; CI = confidence interval.

3.8. Pooled Logistic Regression: Formal Test of Age–Cohort Interaction in Co-infection Risk

To determine whether the contrasting age patterns of co-infection observed in the PE and OPD cohorts (monotonic increase versus U-shaped distribution, respectively; Section 3.3) represented a statistically significant effect modification rather than a chance difference, a pooled logistic regression model with age-group × cohort interaction terms was fitted to the combined HPV-positive dataset. The analytic sample comprised 8,549 HPV-positive women aged 21–82 years (PE: n = 1,963, events = 434; OPD: n = 6,586, events = 494) after restriction to the overlapping age range of both cohorts.

Cohort effect. In the model without interaction terms (Model B), attendance at an OPD clinic was independently associated with markedly lower odds of co-infection compared with PE attendance (OR 0.284, 95% CI 0.246–0.328; $p < 0.001$), after adjustment for age group. This three-fold reduction in co-infection odds associated with OPD attendance persisted and remained quantitatively similar in the fully-specified interaction model (OR 0.271, 95% CI 0.209–0.352; $p < 0.001$), confirming the co-infection

prevalence divergence observed in the descriptive analysis. The full Model B coefficient estimates are presented in Panel A of Table 4.

Likelihood ratio test for interaction. The model including all five age-group × cohort interaction terms (Model C) fit the data significantly better than the model with main effects only (Model B): LRT $\chi^2(5) = 17.62$, $p = 0.0035$. This result formally confirms that the relationship between age group and HPV co-infection risk differs significantly between the PE and OPD cohorts — that is, cohort membership is a statistically significant modifier of the age–co-infection association. Model C demonstrated adequate fit (Hosmer–Lemeshow $p = 1.000$; McFadden $R^2 = 0.061$).

Marginal age-group ORs within each cohort. The complete interaction model coefficients are presented in Panel B of Table 4, and the cohort-stratified marginal ORs are visualized in Figure 7. Within the PE cohort, co-infection risk increased monotonically from the reference group, with significantly elevated ORs in women aged <25 years (OR 5.33, 95% CI 1.41–20.15; $p = 0.014$) and ≥65 years (OR 4.08, 95% CI 2.20–7.56; $p < 0.001$), and a significant intermediate elevation at 55–64 years (OR 1.52, 95% CI 1.07–2.16; $p = 0.019$). Within the OPD cohort, the pattern diverged substantially: the 55–64 year group carried the highest co-infection odds (OR 2.53, 95% CI 1.90–3.39; $p < 0.001$), consistent with and strengthening the finding from the cohort-specific analysis (adjusted OR 4.29 reported in Section 3.7). Younger OPD women also showed elevated co-infection odds relative to the 35–44 year reference (<25 years: OR 1.75, 95% CI 1.12–2.74, $p = 0.014$; 25–34 years: OR 1.35, 95% CI 1.05–1.75, $p = 0.020$), consistent with the U-shaped age pattern observed descriptively. Notably, the ≥65 year group, which showed the highest PE co-infection odds (OR 4.08), was not significantly associated with co-infection in the OPD cohort (OR 1.29, 95% CI 0.61–2.70; $p = 0.506$), and the divergence between these two estimates is visually striking in Figure 6 — the PE and OPD confidence intervals at ≥65 years are almost entirely non-overlapping, with the PE estimate sitting well above the null while the OPD estimate straddles it. The ≥65 × cohort interaction term confirmed this difference (OR 0.32, $p = 0.019$).

Sensitivity analysis. Repeating the analysis using the full OPD age range (15–90 years; n = 8,660; events = 939) produced substantively identical results, with the LRT for interaction remaining significant, confirming that the interaction findings are not an artefact of age range restriction.

Table 4. Pooled binary logistic regression of HPV co-infection among HPV-positive women (age 21–82 years; n = 8,549): main-effects model (Model B), interaction model (Model C), and cohort-stratified marginal ORs derived from Model C.

Panel A: Model B — Age group + Cohort (main effects; no interaction)

Variable	OR	95% CI	p-value
Cohort (OPD vs PE)	0.284	0.246–0.328	<0.001**
Age group (ref: 35–44 years)			
<25	1.891	1.258–2.841	0.002*
25–34	1.262	1.029–1.548	0.025*
45–54	1.175	0.981–1.407	0.081
55–64	2.038	1.626–2.555	<0.001**
≥65	2.380	1.544–3.668	<0.001**

McFadden $R^2 = 0.058$; Hosmer–Lemeshow $\chi^2(8) = 3.32$, $p = 0.913$

Panel B: Model C — Age group + Cohort + Age × Cohort interaction

Variable	OR	95% CI	p-value
Cohort (OPD vs PE)	0.271	0.209–0.352	<0.001**
Age group (ref: 35–44 years, within PE cohort)			
<25	5.333	1.411–20.148	0.014*
25–34	1.103	0.782–1.557	0.576
45–54	1.211	0.932–1.575	0.152
55–64	1.520	1.072–2.156	0.019*
≥65	4.081	2.203–7.558	<0.001**
Age × Cohort interaction terms			
<25 × Cohort	0.328	0.081–1.335	0.120
25–34 × Cohort	1.227	0.799–1.883	0.350
45–54 × Cohort	0.932	0.648–1.341	0.704
55–64 × Cohort	1.666	1.058–2.624	0.028*
≥65 × Cohort	0.315	0.120–0.826	0.019*

McFadden $R^2 = 0.061$; Hosmer–Lemeshow $\chi^2(8) = 0.000$, $p = 1.000$ **LRT Model C vs Model B: $\chi^2(5) = 17.62$, $p = 0.0035$**

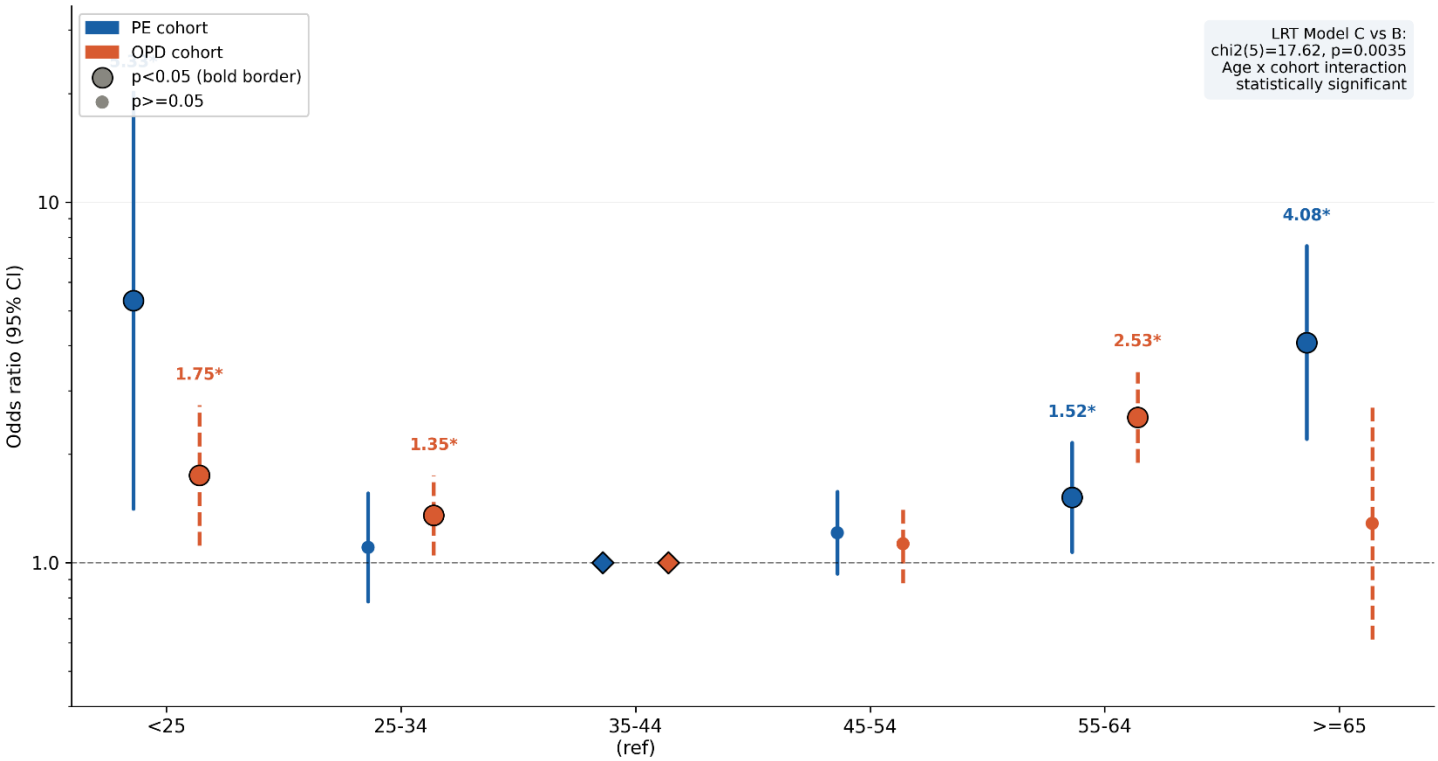


Figure 7. Cohort-stratified marginal odds ratios for HPV co-infection derived from the pooled interaction model (Model C). Points represent marginal odds ratios for each age group relative to the 35–44 year reference (shown as diamonds at OR = 1.0), derived from the pooled binary logistic regression interaction model. Blue circles = Physical Examination (PE) cohort; orange circles = Outpatient Department (OPD) cohort. Vertical lines represent 95% confidence intervals. The y-axis is on a log scale. Bold-bordered circles denote $p < 0.05$; faint circles denote $p \geq 0.05$. Annotated values indicate statistically significant ORs. The inset box reports the likelihood ratio test result for the age × cohort interaction (LRT $\chi^2(5) = 17.62$, $p = 0.0035$), confirming that the age–co-infection relationship differs significantly between cohorts. OR = odds ratio; CI = confidence interval.

4. Discussion

This study provides a comprehensive comparative analysis of HPV co-infection patterns across two clinically distinct populations; asymptomatic women attending routine cervical cancer screening and symptomatic women presenting to outpatient gynecology services, drawing on a large, genotype-resolved dataset of nearly 38,000 women from Taizhou, China. The most striking finding is the paradoxically higher co-infection rate in the asymptomatic PE cohort (22.1%) compared with the symptomatic OPD cohort (7.5%), despite the latter carrying a significantly higher overall HPV positivity rate (24.0% vs. 19.5%). This apparent contradiction resolves when the denominator structure is considered: the OPD cohort's substantially larger pool of HPV-positive women dilutes the co-infection proportion, while the clinical dynamics of symptomatic presentation favor dominant single high-risk infections (particularly HPV16, which ranked second in the OPD cohort but only fourth in the PE cohort) as the primary drivers of gynecological consultation. This pattern is consistent with observations from other high-volume Asian screening programmes, where clinical-encounter populations show enrichment for the most oncogenic

genotypes relative to population-based screening cohorts [24,25]. The ascendancy of HPV16 in the OPD setting carries direct triage implications: current Chinese cervical cancer screening guidelines recommend immediate colposcopy referral for HPV16/18-positive women [26], and our data suggest that this threshold is well-calibrated to the genotypic reality of symptomatic outpatient populations, where HPV16 burden is disproportionately concentrated.

The age patterning of co-infection represents a second major finding of this analysis, and one with important mechanistic implications. The PE cohort exhibited a monotonic increase in co-infection prevalence with advancing age — confirmed by the Cochran–Armitage test ($p = 0.002$) and Spearman's correlation ($\rho = 0.070$) — consistent with the hypothesis of cumulative multi-genotype acquisition over the lifespan combined with progressive impairment of immunological clearance. This pattern mirrors findings from population-based screening cohorts in other East Asian settings, where age-related immune senescence has been proposed as the primary driver of multi-genotype persistence in older women [27,28]. The OPD cohort, by contrast, showed a U-shaped age distribution, with co-infection rates elevated in both the youngest (<25 years) and oldest (≥55 years) women and a nadir at 35–44 years. The early peak in younger OPD attendees most plausibly reflects recent multi-partner exposure and acute simultaneous acquisition of multiple genotypes, consistent with the higher sexual behavior risk profiles associated with symptomatic gynecological presentations in younger women [29]. The late peak — and particularly the striking adjusted OR of 4.29 (95% CI 2.16–8.52) in women aged 55–64 years attending OPD — is less readily explained by behavioural factors and more likely reflects immunosenescence-driven persistence of multiple previously acquired genotypes, possibly compounded by the hormonal immune modulation associated with the peri-menopausal transition [18]. The formal pooled interaction analysis confirmed that these divergent age patterns are not a chance difference but a statistically significant effect modification by clinical setting (LRT $\chi^2(5) = 17.62$, $p = 0.0035$), establishing that the age–co-infection relationship is fundamentally different between screening and clinical-encounter populations.

The genotype co-occurrence analysis revealed a pervasive pattern of competitive exclusion (co-occurrence avoidance) between the most prevalent hrHPV types in both cohorts, a finding that was substantially more pronounced in the OPD population. In the PE cohort, the most frequently observed pairs, including HPV52+HPV58 (O/E = 0.48) and HPV16+HPV52 (O/E = 0.46), occurred at less than half the frequency expected under the independence assumption. In the OPD cohort, this exclusion intensified dramatically, with HPV16+HPV52 showing an O/E of just 0.14 — the most extreme avoidance signal in the entire dataset. Competitive exclusion between HPV genotypes has been documented in multiple epidemiological studies and attributed to several non-mutually exclusive mechanisms, including type-specific immune cross-reactivity, competition for susceptible epithelial cell populations, and interference at the level of viral replication [9,30]. The amplification of this signal in the OPD cohort is consistent with the hypothesis that dominant high-risk genotypes exert stronger competitive suppression of co-colonization in a clinically enriched population, though the cross-sectional design of this study precludes causal inference about the directionality or mechanism of this relationship. Importantly, the O/E ratio heatmap (Figure 4) visualizes this landscape across all 105 possible genotype pairs simultaneously, revealing that competitive exclusion

is the dominant biological signal in the hrHPV–hrHPV interaction space in both cohorts, while isolated instances of positive affinity are confined largely to lower-prevalence lrHPV type pairs.

The hrHPV co-infection analysis adds a clinically critical dimension to these findings. Among co-infected women, the OPD cohort showed a markedly higher proportion of hrHPV+hrHPV dual infections (64.6%) compared with the PE cohort (43.0%), despite the OPD cohort's lower overall co-infection rate. This means that when OPD women are co-infected, they are disproportionately co-infected with two or more definite carcinogenic genotypes — a combination that may confer greater cumulative oncogenic risk than single high-risk infection, as suggested by studies linking dual hrHPV infection to higher rates of high-grade cervical intraepithelial neoplasia [31,32]. The dominance of HPV52 across both cohorts; ranking first in both, accounting for 17.4% and 18.5% of HPV-positive women in the PE and OPD cohorts respectively, reinforces the well-established East Asian epidemiological signature of HPV52/58 predominance and has direct implications for vaccine policy [33,34]. The bivalent and quadrivalent vaccines, which target only HPV16 and HPV18 (with or without HPV6/11), would provide no direct protection against HPV52 or HPV58. The nonavalent vaccine, which additionally covers HPV31, 33, 45, 52, and 58, would therefore confer substantially broader protection in this population [35]. Our data provide further impetus for accelerating nonavalent vaccine uptake in China, where the national immunization programme has expanded HPV vaccination but coverage of the nonavalent formulation remains heterogeneous across regions [36].

Several limitations of this study warrant consideration. First, the cross-sectional design precludes causal inference and prevents assessment of temporal co-infection dynamics — it is not possible to determine whether competitive exclusion reflects true biological interference or differential clearance rates of co-infecting genotypes over time. Second, the HybriMax® genotyping panel detects 21 genotypes, meaning that infections with non-panel genotypes are undetected, and true co-infection prevalence may be underestimated in both cohorts. Third, the absence of cytological, colposcopic, or histological data prevents direct assessment of the clinical consequences of co-infection patterns identified here, and future linkage of genotyping data to cervical disease endpoints would substantially strengthen the clinical translation of these findings. Fourth, residual confounding by sexual behavior, parity, contraceptive use, and immunosuppressive comorbidity cannot be excluded. Notwithstanding these limitations, the scale of the dataset, the systematic comparative design across two clinically distinct populations, the formal interaction-testing framework, and the granular 15-genotype O/E analysis represent meaningful contributions to the HPV co-infection literature. The finding that women aged 55–64 years attending outpatient gynecology services carry a fourfold elevation in multi-genotype co-infection odds is the single most actionable clinical implication of this work, and supports consideration of extended genotyping panels and expedited colposcopy referral pathways for peri- and post-menopausal women presenting to gynecology outpatient services in high-burden settings.

Conclusion

In summary, this large-scale comparative analysis demonstrates that HPV co-infection prevalence, genotype pairing dynamics, and age-patterning differ fundamentally between asymptomatic cervical cancer screening populations and symptomatic outpatient gynecology populations — differences that are not merely quantitative but qualitatively distinct, as confirmed by formal statistical interaction testing. The paradoxically higher co-infection rate in the asymptomatic PE cohort, the U-shaped age–co-infection relationship unique to the OPD setting, the fourfold elevation in co-infection odds among OPD women aged 55–64 years, the enrichment of HPV16 and hrHPV+hrHPV dual infections in the clinical-encounter population, and the intensification of competitive exclusion between dominant oncogenic genotypes in the outpatient setting collectively underscore that clinical context is a fundamental modifier of HPV co-infection epidemiology and must be explicitly accounted for in the design and interpretation of prevalence surveys, triage algorithms, and vaccine impact assessments.

The dominance of HPV52 across both cohorts reaffirms the priority of nonavalent over lower-valency vaccine formulations in this population, while the disproportionate co-infection burden in peri- and post-menopausal OPD attendees identifies a high-risk subgroup that warrants consideration for extended genotyping and expedited colposcopy referral in routine clinical practice. Future prospective studies linking genotype-resolved co-infection data to longitudinal cervical disease outcomes across both screening and clinical settings will be essential to translate these epidemiological observations into evidence-based clinical guidance.

Acknowledgement

Not applicable

Author contribution

Conceptualization, YZ; methodology; YZ, and ET; software, YZ and ET; validation, CCB; formal analysis, YZ and ET; resources CCB; data curation, YZ.; writing—original draft preparation YZ; writing—review and editing, ET, CCB; supervision, CCB; project administration CCB; All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Ethics approval and written informed consent for the original data collection were obtained by the primary investigators from the Institutional Medical Ethics Review Board of Taizhou Hospital in Zhejiang Province. Our secondary use of the anonymized dataset did not require additional ethics review.

Consent to participate: Not applicable

Consent to publish: Not applicable

Data Availability Statement: All data supporting the conclusions of this article are online at https://doi.org/10.5061/dryad.c7h3c

Conflicts of Interest: The authors declare no conflicts of interest

Funding: This research received no external funding

Clinical trial number: Not applicable

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63. doi:10.3322/caac.21834 PubMed PMID: 38572751.

2. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol. 1999 Sep;189(1):12–9. doi:10.1002/(SICI)1096-9896(199909)189:1%3C12::AID-PATH431%3E3.0.CO;2-F PubMed PMID: 10451482.

3. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. The Lancet. 2007 Sep 8;370(9590):890–907. doi:10.1016/S0140-6736(07)61416-0 PubMed PMID: 17826171.

4. Han J, Zhang L, Chen Y, Zhang Y, Wang L, Cai R, et al. Global HPV vaccination programs and coverage rates: a systematic review. EClinicalMedicine. 2025 Jun;84:103290. doi:10.1016/j.eclinm.2025.103290 PubMed PMID: 40547442; PubMed Central PMCID: PMC12179740.

5. Wu S, Jiao J, Yue X, Wang Y. Cervical cancer incidence, mortality, and burden in China: a time-trend analysis and comparison with England and India based on the global burden of disease study 2019. Front Public Health. 2024;12:1358433. doi:10.3389/fpubh.2024.1358433 PubMed PMID: 38510348; PubMed Central PMCID: PMC10951371.

6. Han S, Lin M, Liu M, Wu S, Guo P, Guo J, et al. Prevalence, trends, and geographic distribution of human papillomavirus infection in Chinese women: a summative analysis of 2,728,321 cases. BMC Med. 2025 Mar 13;23:158. doi:10.1186/s12916-025-03975-6 PubMed PMID: 40082952; PubMed Central PMCID: PMC11907810.

7. Mo B, Ye Y, Yu M, Tong X, Cao H, Du C, et al. Prevalence and genotype distribution of HPV combined with cervical pathological results in women from Sichuan, China: A cross-sectional study based on post-vaccination period 2019 to 2023. Cancer Med. 2024 Aug;13(16):e70148. doi:10.1002/cam4.70148 PubMed PMID: 39189602; PubMed Central PMCID: PMC11348228.

8. Li M, Wei L, Sui L, Ma D, Kong B, Wu X, et al. Guidelines for cervical cancer screening in China. Gynecology and Obstetrics Clinical Medicine. 2023 Dec;3(4):189–94. doi:10.1016/j.gocm.2023.10.005

9. Vaccarella S, Franceschi S, Snijders PJF, Herrero R, Meijer CJLM, Plummer M, et al. Concurrent Infection with Multiple Human Papillomavirus Types: Pooled Analysis of the IARC HPV Prevalence Surveys. Cancer Epidemiol Biomarkers Prev. 2010 Feb 7;19(2):503–10. doi:10.1158/1055-9965.EPI-09-0983

10. Ma Y, Xia X, Zheng W, Dai Y, Zhuang X. HPV prevalence and genotype distribution among women in eastern China during the Covid-19 pandemic. Human Vaccines & Immunotherapeutics. 2023 Jan 2;19(1):2212571. doi:10.1080/21645515.2023.2212571 PubMed PMID: 37226673.

11. Tang X, Zhang H, Wang T, Jiang W, Jones TE, He Y, et al. Single and Multiple High-Risk Human Papillomavirus Infections in Histopathologically Confirmed Cervical Squamous Lesions: Incidences, Distribution, and Associated Detection Rates for Precancerous and Cancerous Lesions. Lab Invest. 2023 Nov;103(11):100234. doi:10.1016/j.labinv.2023.100234 PubMed PMID: 37574009.

12. Cassani C, Dominoni M, Pasquali MF, Gardella B, Spinillo A. The role of multiple high-risk human papillomavirus infection on the persistence recurrence of high-grade cervical lesions after standard treatment: A systematic review and a meta-analysis. Acta Obstet Gynecol Scand. 2024 Jun;103(6):1028–35. doi:10.1111/aogs.14827 PubMed PMID: 38477097; PubMed Central PMCID: PMC11103134.

13. Luo Q, Zeng X, Luo H, Pan L, Huang Y, Zhang H, et al. Epidemiologic characteristics of high-risk HPV and the correlation between multiple infections and cervical lesions. BMC Infect Dis. 2023 Oct 7;23(1):667. doi:10.1186/s12879-023-08634-w PubMed PMID: 37805467; PubMed Central PMCID: PMC10560423.

14. Kim JM, Song HS, Hwang J, Kim JK. The Prevalence of Multi-Type Infections Among Human Papillomavirus Types in Korean Women. Pathogens. 2025 Apr;14(4):369. doi:10.3390/pathogens14040369

15. Zheng ZM, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. Front Biosci. 2006 Sep 1;11:2286–302. doi:10.2741/1971 PubMed PMID: 16720315; PubMed Central PMCID: PMC1472295.

16. Veldhuijzen NJ, Snijders PJ, Reiss P, Meijer CJ, van de Wijgert JH. Factors affecting transmission of mucosal human papillomavirus. Lancet Infect Dis. 2010 Dec;10(12):862–74. doi:10.1016/S1473-3099(10)70190-0 PubMed PMID: 21075056.

17. Man I, Vänskä S, Lehtinen M, Bogaards JA. Human Papillomavirus Genotype Replacement: Still Too Early to Tell? J Infect Dis. 2021 Aug 2;224(3):481–91. doi:10.1093/infdis/jiaa032 PubMed PMID: 31985011; PubMed Central PMCID: PMC8328199.

18. Gravitt PE. The known unknowns of HPV natural history. J Clin Invest. 2011 Dec;121(12):4593–9. doi:10.1172/JCI57149 PubMed PMID: 22133884; PubMed Central PMCID: PMC3225991.

19. Schiffman M, Doorbar J, Wentzensen N, de Sanjosé S, Fakhry C, Monk BJ, et al. Carcinogenic human papillomavirus infection. Nat Rev Dis Primers. 2016 Dec 1;2:16086. doi:10.1038/nrdp.2016.86 PubMed PMID: 27905473.

20. Wentzensen N, Schiffman M, Palmer T, Arbyn M. Triage of HPV positive women in cervical cancer screening. J Clin Virol. 2016 Mar;76 Suppl 1(Suppl 1):S49–55. doi:10.1016/j.jcv.2015.11.015 PubMed PMID: 26643050; PubMed Central PMCID: PMC4789103.

21. Luo Q, Zeng X, Luo H, Pan L, Huang Y, Zhang H, et al. Epidemiologic characteristics of high-risk HPV and the correlation between multiple infections and cervical lesions. BMC Infect Dis. 2023 Oct 7;23(1):667. doi:10.1186/s12879-023-08634-w PubMed PMID: 37805467; PubMed Central PMCID: PMC10560423.

22. Hou J, Zeng M, Liu C, Xie B, Li Y, Wu L, et al. Cervical HPV infection in Yueyang, China: a cross-sectional study of 125,604 women from 2019 to 2022. Front Public Health. 2023 Aug 3;11:1210253. doi:10.3389/fpubh.2023.1210253 PubMed PMID: 37601194; PubMed Central PMCID: PMC10435747.

23. Xu HH, Lin A, Chen YH, Dong SS, Shi WW, Yu JZ, et al. Prevalence characteristics of cervical human papillomavirus (HPV) genotypes in the Taizhou area, China: a cross-sectional study of 37 967 women from the general population. BMJ Open. 2017 Jun 9;7(6):e014135. doi:10.1136/bmjopen-2016-014135 PubMed PMID: 28601819; PubMed Central PMCID: PMC5577888.

24. Zhao FH, Lewkowitz AK, Hu SY, Chen F, Li LY, Zhang QM, et al. Prevalence of Human Papillomavirus and Cervical Intraepithelial Neoplasia in China: A pooled analysis of 17 Population-based Studies. Int J Cancer. 2012 Dec 15;131(12):2929–38. doi:10.1002/ijc.27571 PubMed PMID: 22488743; PubMed Central PMCID: PMC3435460.

25. Zhang W, Guo N, Li B, Shang E, Wang J, Zhang M, et al. Prevalence and genotype distribution of human papillomavirus infections in Beijing, China between 2016 and 2020. Virol J. 2023 Jan 18;20:11. doi:10.1186/s12985-023-01959-7 PubMed PMID: 36653807; PubMed Central PMCID: PMC9847084.

26. Li M, Wei L, Sui L, Ma D, Kong B, Wu X, et al. Guidelines for cervical cancer screening in China. Gynecology and Obstetrics Clinical Medicine. 2023 Dec;3(4):189–94. doi:10.1016/j.gocm.2023.10.005

27. Chung S, Shin S, Yoon JH, Roh EY, Seoung SJ, Kim GP, et al. Prevalence and Genotype of Human Papillomavirus Infection and Risk of Cervical Dysplasia among Asymptomatic Korean Women. Ann Clin Microbiol. 2013 Jun 5;16(2):87–91. doi:10.5145/ACM.2013.16.2.87

28. Zheng ZM, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. Front Biosci. 2006 Sep 1;11:2286–302. doi:10.2741/1971 PubMed PMID: 16720315; PubMed Central PMCID: PMC1472295.

29. Veldhuijzen NJ, Snijders PJ, Reiss P, Meijer CJ, van de Wijgert JH. Factors affecting transmission of mucosal human papillomavirus. Lancet Infect Dis. 2010 Dec;10(12):862–74. doi:10.1016/S1473-3099(10)70190-0 PubMed PMID: 21075056.

30. Pintos J, Black MJ, Sadeghi N, Ghadirian P, Zeitouni AG, Viscidi RP, et al. Human papillomavirus infection and oral cancer: a case-control study in Montreal, Canada. Oral Oncol. 2008 Mar;44(3):242–50. doi:10.1016/j.oraloncology.2007.02.005 PubMed PMID: 17467327.

31. Trottier H, Mahmud S, Costa MC, Sobrinho JP, Duarte-Franco E, Rohan TE, et al. Human papillomavirus infections with multiple types and risk of cervical neoplasia. Cancer Epidemiol Biomarkers Prev. 2006 Jul;15(7):1274–80. doi:10.1158/1055-9965.EPI-06-0129 PubMed PMID: 16835323.

32. Chaturvedi AK, Katki HA, Hildesheim A, Rodríguez AC, Quint W, Schiffman M, et al. Human papillomavirus infection with multiple types: pattern of coinfection and risk of cervical disease. J Infect Dis. 2011 Apr 1;203(7):910–20. doi:10.1093/infdis/jiq139 PubMed PMID: 21402543; PubMed Central PMCID: PMC3068034.

33. Li J, Huang R, Schmidt JE, Qiao YL. Epidemiological features of Human Papillomavirus (HPV) infection among women living in Mainland China. Asian Pac J Cancer Prev. 2013;14(7):4015–23. doi:10.7314/apjcp.2013.14.7.4015 PubMed PMID: 23991946.

34. Bao YP, Li N, Smith JS, Qiao YL, ACCPAB members. Human papillomavirus type distribution in women from Asia: a meta-analysis. Int J Gynecol Cancer. 2008;18(1):71–9. doi:10.1111/j.1525-1438.2007.00959.x PubMed PMID: 17466054.

35. Joura EA, Giuliano AR, Iversen OE, Bouchard C, Mao C, Mehlsen J, et al. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med. 2015 Feb 19;372(8):711–23. doi:10.1056/NEJMoa1405044 PubMed PMID: 25693011.

36. Guo F, Cofie LE, Berenson AB. Cervical Cancer Incidence in Young U.S. Females After Human Papillomavirus Vaccine Introduction. Am J Prev Med. 2018 Aug;55(2):197–204. doi:10.1016/j.amepre.2018.03.013 PubMed PMID: 29859731; PubMed Central PMCID: PMC6054889.