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.

Immediate Risk of Cervical Precancer and Primary Triage Strategies for Women with Non- 16/18 High-Risk HPV Infection and ASC-US/LSIL Cytology: A Real-World Cross-Sectional Study

Jun Ma
Ningbo University Affiliated Women and Children's Hospital
Jue Zhu
Ningbo University Affiliated Women and Children's Hospital
Linghui Weng
Ningbo University Affiliated Women and Children's Hospital
Yichen Chen
Ningbo University Affiliated Women and Children's Hospital
Liping Chen
Ningbo University Affiliated Women and Children's Hospital
Liang Chen
Ningbo University Affiliated Women and Children's Hospital
Xiao Li
Women's Hospital, School of Medicine, Zhejiang University
Jing Zhang
zhangjyy1978@163.com

Ningbo University Affiliated Women and Children's Hospital

Research Article

Keywords: High-risk HPV (hrHPV), Extended genotyping, Low-grade squamous intraepithelial lesion (LSIL), Atypical squamous cells of undetermined significance (ASC-US)

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11677277/v1

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

Additional Declarations: No competing interests reported.

Abstract

Objective

We evaluated the immediate risk of cervical intraepithelial neoplasia grade 2 or worse (CIN2+) conferred by non-16/18 high-risk HPV (hrHPV) genotypes in women with atypical squamous cells of undetermined significance (ASC-US) and low-grade squamous intraepithelial lesion (LSIL) cytology, and developed a risk-based primary triage strategy. We also characterized the regional epidemiological features of HPV infection in Ningbo, China, from 2019 to 2024.

Methods

Based on defined eligibility criteria, this study enrolled 511,133 women participating in cervical cancer screening in Ningbo from 2019 to 2024. Colposcopy and histopathological biopsy were performed based on HPV and liquid-based cytology (TCT) findings. We used lowest-attribution (minimum), any-attribution, and hierarchical-attribution models to evaluate the distribution of non-16/18 hrHPV genotypes and their immediate risks of progressing to CIN2 + under ASC-US and LSIL cytology. Finally, we formulated a risk-stratified clinical triage framework by integrating recent evidence on extended HPV genotyping.

Results

Of 511,133 women screened, 50,320 (9.84%) were HPV-positive, with HPV52, 58, and 16 identified as the predominant genotypes. Non-16/18 hrHPV accounted for 84.18% (42,357/50,320) of all HPV-positive cases. Across both ASC-US and LSIL cytology, HPV31 conferred the highest immediate risk of CIN2 + among all non-16/18 genotypes.

Conclusion

Although non-16/18 hrHPV genotypes predominate among HPV-positive women, their immediate risks of CIN2 + are highly heterogeneous. A uniform management protocol for managing non-16/18 hrHPV infections wastes medical resources. We propose a three-tiered, extended genotyping-based triage strategy for women with ASC-US/LSIL cytology: direct colposcopy for high-risk genotypes (HPV31, 33, 52, and 58); conservative follow-up for low-risk genotypes (HPV39, 56, 66, and 68); and secondary triage using biomarkers (e.g., DNA methylation, p16/Ki-67 dual staining) or cytology for intermediate-risk genotypes. This framework is designed for initial, immediate triage and does not replace long-term longitudinal surveillance for specific genotypes.

Key Messages

1. Non-16/18 hrHPV genotypes exhibit profound pathogenic heterogeneity in women with ASC-US/LSIL cytology, indicating that a uniform management protocol leads to the massive misallocation of medical resources.
2. We propose a refined three-tier triage strategy where extremely high-risk genotypes (HPV31, 33, 52, and 58) warrant immediate colposcopy, while low-risk genotypes (HPV39, 56, 66, and 68) can be safely managed with conservative follow-up.
3. Implementing this extended genotyping-based framework successfully improves screening specificity by 40.11% and optimizes the health-economic efficiency of cervical cancer prevention.

1. Introduction

Cervical cancer is one of the few malignancies that can be effectively prevented through routine screening. Persistent high-risk human papillomavirus (hrHPV) infection is its primary cause. Current screening paradigms are predominantly HPV-based, guiding downstream triage and clinical intervention according to specific viral genotypes and concurrent cytology. Because HPV16 and 18 cause roughly 70% of cervical cancers and drive high-grade lesion development [1], major guidelines, including the ASCCP consensus, recommend immediate colposcopy without secondary testing [2]. Conversely, existing guidelines generally endorse a uniform cytology-based triage pathway for non-16/18 hrHPV infections. This population comprises 70%-85% of all HPV-positive individuals [3–4]. Given the broad spectrum of non-16/18 genotypes (including HPV31, 33, 45, 52, and 58) and their widely variable oncogenic risks, this undifferentiated approach leads to excessive colposcopic examinations and unwarranted psychological distress [3, 5]. The World Health Organization (WHO) recently released its updated guidance on cervical pre-cancer screening and treatment, reinforcing the clinical importance of genotype-specific risk assessment [6]. Therefore, precise evaluation of the immediate risk of cervical intraepithelial neoplasia grade 2 or worse (CIN2+) across extended non-16/18 genotypes under abnormal cytology and the formulation of tailored triage algorithms have emerged as critical challenges in contemporary cervical cancer prevention.

In this multicenter, real-world cross-sectional study based on a massive screening cohort in Ningbo, China, we systematically evaluated local HPV epidemiology and determined the genotype-specific immediate and independent risks of CIN2+ (and invasive cancer) conferred by non-16/18 hrHPV in women with ASC-US/LSIL cytology, then we developed and validated a refined triage strategy. Designed specifically as an initial, immediate, triage framework, this strategy aims to optimize current clinical pathways and alleviate the burden on colposcopy services; however, it is not intended to substitute for long-term longitudinal surveillance of genotypes with slower oncogenic potential.

2. Materials and Methods

2.1. Study population

In this multicenter retrospective study, we retrieved community-based cervical cancer screening records collected from Ningbo, China (including Xiangshan County, Yuyao City, Yinzhou District, Haishu District, Ninghai County, and Fenghua District). The cervical cancer screening program was implemented in accordance with the Cervical Cancer Screening Program Guidelines issued by the National Health Commission of the People's Republic of China, primarily targeting female permanent residents aged 35–64 years. The inclusion criteria were as follows: (1) voluntary participation in HPV testing; (2) registered residence in Ningbo; and (3) testing via extended HPV genotyping platforms. Exclusion criteria were defined as: (1) a prior history of cervical cancer, total hysterectomy, or colpocleisis; (2) pregnancy or being within six weeks postpartum or post-abortion; and (3) incomplete clinical records or unknown definitive follow-up outcomes. These criteria ensured cohort consistency and the complete tracking of the primary outcome-CIN2+. This study protocol was approved by the Medical Ethics Committee of Ningbo University Affiliated Women and Children's Hospital (Approval No.: NBFE-2026-KY-005). Given the retrospective nature of this

study, which utilized de-identified historical screening records, the requirement for obtaining written informed consent from individual patients was waived by the Medical Ethics Committee of Ningbo University Affiliated Women and Children's Hospital.

2.2. Data collection and clinical management principles

Clinical baseline records—including age, HPV genotyping, liquid-based cytology (TCT), and histological diagnoses-were systematically retrieved. Cervical specimens were obtained following standardized clinical protocols: trained clinicians collected cervical cells using dedicated cytobrushes, transferred them into vials containing cell-preservative medium, and immediately transported them for laboratory testing in accordance with biosafety guidelines. Colposcopy was performed by experienced colposcopic impression, during which targeted biopsies, random biopsies, or endocervical curettage (ECC) were conducted as clinically indicated. Throughout the retrospective study period, clinical triage adhered to the following principles (Fig. 1): women testing positive for HPV16/18 were referred directly for colposcopy and histopathological biopsy without prior cytology triage. Conversely, women harboring non-16/18 hrHPV genotypes underwent reflexive TCT triage, and only those with abnormal cytological findings were referred for subsequent colposcopic evaluation and histopathological assessment.

HPV: human papillomavirus; TCT: liquid-based thin-layer cytology; ASC-US: atypical squamous cells of undetermined significance; LSIL: low-grade squamous intraepithelial lesion; NILM: negative for intraepithelial lesion or malignancy.

2.3. HPV genotyping assays and histopathological criteria

HPV genotyping was performed using four widely adopted commercial real-time fluorescence PCR assays: Dian Diagnostics (Hangzhou, China), MedicalSystem Biotechnology (Ningbo, China), BGI Genomics (Shanghai, China), and Liferiver Bio-Tech (Shanghai, China). All four assays individually covered 17 high-risk HPV (hrHPV) genotypes (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, and 82) and four low-risk HPV (lrHPV) genotypes (HPV6, 11, 42, and 81).

Histopathology serves as the gold standard for diagnosing cervical lesions. All histopathological specimens were independently reviewed and reported by two senior pathologists, and discrepancies were resolved by consensus (or consultation with a third expert). Diagnoses were established according to the 2020 World Health Organization (WHO) Classification of Female Genital Tumours, categorized as: normal, inflammation, cervical intraepithelial neoplasia grade 1 (CIN1), 2 (CIN2), and 3 (CIN3), adenocarcinoma in situ (AIS), and invasive cervical cancer (including squamous cell carcinoma and adenocarcinoma) [7]. In this study, cervical intraepithelial neoplasia grade 2 or worse (CIN2+) was defined as the primary clinical endpoint, encompassing CIN2, CIN3, AIS, and cervical cancer. High-grade vaginal intraepithelial neoplasia (VaIN 2 or 3) was categorized as CIN2 or CIN3, respectively, for analytical purposes.

2.4. Statistical analysis

For age-stratified analyses, participants were categorized into six age brackets: ≤39, 40–44, 45–49, 50–54, 55–59, and ≥ 60 years. To evaluate the epidemiological distribution of individual HPV genotypes, both any-type (Any.) and minimum (Min.) estimation approaches were applied. When evaluating genotype-specific

distributions and immediate risks of CIN2 + in women with ASC-US or LSIL, we adopted three attribution models as described by Li et al. [8–9]: (1) Minimum estimate (Min.): restricted to women harboring single-genotype HPV infections; (2) Any-type estimate (Any.): accounting for all detections of a given genotype, where multiple infections were counted repeatedly across each detected type; and (3) Hierarchical-attribution estimate (Hier.): hierarchically attributing co-infections to the single genotype conferring the highest immediate risk of CIN2 + determined by the Any. method in the overall screening cohort. When applying the Hier. algorithm to resolve multiple infections, establishing a predefined pathogenic hierarchy among extended genotypes was required. To minimize stochastic bias arising from small sample sizes of rare genotypes within the ASC-US and LSIL subsets, a global baseline priority ranking was formulated based on the immediate risk of CIN2 + under the Any. algorithm across the entire screening cohort (N = 511,133). This pathogen-ranking hierarchy, robustly derived from the large population cohort, was subsequently applied to the targeted ASC-US and LSIL subgroups to disentangle co-infections and quantify the independent oncogenic risk of each non-16/18 genotype under specific cytological backgrounds.

After data curation and verification, statistical analyses were conducted using IBM SPSS Statistics (Version 31.0; IBM Corp., Armonk, NY, USA), with graphical visualization performed using GraphPad Prism (Version 10.0; GraphPad Software,San Diego, CA, USA) and R software (Version 4.5.3; R Foundation for Statistical Computing, Vienna, Austria). Categorical variables (e.g., HPV infection rates and histopathological grade distributions) were presented as counts (*n*) and percentages (%). Differences in rates or proportions between groups were compared using the Pearson chi-square (χ^2) test, or Fisher's exact test when expected cell frequencies were less than 5. All statistical tests were two-sided, with $P < 0.05$ considered statistically significant.

To determine triage priorities among individual extended HPV genotypes within the positive cohort, the overall average pathogenicity of the non-16/18 hrHPV-positive population was established as the baseline reference for odds ratio (OR) evaluations. To avoid overlap bias in statistical calculations, when calculating the OR for any specific target genotype (e.g., HPV52), the reference group was strictly defined as all remaining non-16/18 hrHPV-positive women after excluding that specific genotype. An OR significantly greater than 1.0 indicated a pathogenic risk significantly above the average cohort level. The clinical performance of the conventional management strategy versus the risk-stratified triage strategy in detecting CIN2 + was evaluated using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the number of colposcopies needed to detect one CIN2 + case (NNC). Furthermore, considering that the screening cohort spanned six years and utilized four mainstream commercial PCR assays, sensitivity analyses using chi-square (χ^2) tests were performed for the predominant pathogenic genotypes (such as HPV52 and HPV58) to assess whether assay-related heterogeneity introduced bias into the estimation of immediate CIN2 + risks.

3. Results

3.1. Baseline characteristics of the study population

3.1.1. Demographic characteristics

A total of 511,133 women undergoing cervical cancer screening in Ningbo between 2019 and 2024 were included in this study. The overall cohort had a median age of 53 years (range: 27–68 years), whereas HPV-positive women were slightly older, with a median age of 55 years (range: 27–65 years) (Supplementary Table A.1). As illustrated in Fig. 2, HPV positivity differed significantly across age groups and increased progressively with age. Specifically, the lowest prevalence was observed in the ≤ 39-year-old group (7.13%), exceeded 10% in the 55-59-year-old group (11.10%), and peaked in the ≥ 60-year-old group (12.79%). Among women participating in cervical cancer screening in Ningbo, perimenopausal and postmenopausal women (aged ≥ 50 years) not only constituted the majority of the screened cohort but also represented the high-prevalence group for HPV infection.

The bars (left y-axis) represent the total number of women screened in each age group, and the line graph (right y-axis) depicts the corresponding HPV positivity rate (%). The data demonstrate a progressive upward trend in HPV positivity with advancing age, peaking in the ≥ 60-year-old group. HPV: human papillomavirus.

As shown in Fig. 3, further longitudinal time-series analysis revealed dynamic temporal trends in overall and age-specific HPV prevalence from 2019 to 2024 (Supplementary Table A.2). Overall HPV positivity peaked in 2019 (11.65%) before declining markedly in 2021 to a six-year nadir of 8.35%—a trough potentially attributable to COVID-19 containment measures and social distancing restrictions that reduced viral exposure. From 2022 to 2024, following the resumption of normal societal activities, overall HPV prevalence rebounded modestly and steadily, reaching 9.64% in 2024 while remaining below pre-pandemic levels. Notably, despite year-to-year fluctuations in overall prevalence, the epidemiological pattern of increasing HPV positivity with advancing age remained highly consistent across all calendar years, with the lowest and highest rates invariably observed in the ≤ 39- and ≥ 60-year-old groups, respectively.

The line graph illustrates the six-year longitudinal changes in HPV infection rates (%) across six age brackets. Differently colored lines represent distinct age groups, with the color gradient (from light yellow to dark red) corresponding to advancing age (from the ≤ 39-year-old group to the ≥ 60-year-old group). Across all age strata, HPV prevalence reaches a trough during 2021–2022 before rebounding steadily. Throughout the six-year study period, HPV prevalence consistently exhibits an age-dependent upward pattern, with the ≥ 60-year-old group (dark red line) persistently bearing the highest infection burden. HPV: human papillomavirus.

3.1.2. Overall HPV positivity and genotype distribution

Among the 511,133 women enrolled, 50,320 tested positive for HPV, corresponding to an overall positivity rate of 9.84%. Single-genotype infections predominated, accounting for 84.36% (42,452/50,320) of positive cases, whereas multiple infections accounted for 15.64% (7,868/50,320). Across all detected HPV genotypes, HPV52 exhibited not only the highest overall prevalence in the screened population (3.31%, 16,935/511,133) but also the largest proportion within the HPV-positive cohort (33.65%, 16,935/50,320), followed by the predominant genotypes HPV58 (15.11%) and HPV16 (10.25%) (Table 1 and Fig. 4).

Table 1

Distribution characteristics of individual HPV genotypes among 511,133 women undergoing cervical cancer screening in Ningbo, China.

HPV genotype	Any.(*n* = 60,045)			Min.(*n* = 42,452)		
	Positive cases(*n*)	Prevalence(%)	Proportion(%)	Positive cases(*n*)	Prevalence(%)	Proportion(%)
HPV16	5,159	1.01	10.25	3,737	0.73	7.43
HPV18	2,965	0.58	5.89	2,142	0.42	4.26
HPV31	2,104	0.41	4.18	1,406	0.28	2.79
HPV33	3,567	0.70	7.09	2,462	0.48	4.89
HPV35	2,353	0.46	4.68	1,497	0.29	2.97
HPV39	2,215	0.43	4.40	1,392	0.27	2.77
HPV45	1,357	0.27	2.70	857	0.17	1.70
HPV51	4,369	0.85	8.68	2,914	0.57	5.79
HPV52	16,935	3.31	33.65	13,465	2.63	26.76
HPV53	89	0.02	0.18	7	0.00	0.01
HPV56	2,819	0.55	5.60	1,735	0.34	3.45
HPV58	7,602	1.49	15.11	5,536	1.08	11.00
HPV59	2,302	0.45	4.57	1,455	0.28	2.89
HPV66	944	0.18	1.88	582	0.11	1.16
HPV68	4,265	0.83	8.48	2,824	0.55	5.61
HPV73	5	0.00	0.01	0	0.00	0.00
HPV82	184	0.04	0.37	97	0.02	0.19
HPV6	496	0.10	0.99	245	0.05	0.49
HPV11	200	0.04	0.40	95	0.02	0.19
HPV42	41	0.01	0.08	1	0.00	0.00
HPV81	79	0.02	0.16	3	0.00	0.01

Note: Any. (any-type estimate, totaling 60,045 genotype detections) includes all cases positive for a given HPV genotype regardless of co-infections; consequently, the sum of proportions across all genotypes exceeds 100%. Min. (minimum estimate, *n* = 42,452) includes only cases with single-genotype infections. Prevalence (%) was calculated as the number of positive cases divided by the total screened population (*N* = 511,133) × 100%. Proportion (%) was calculated as the number of positive cases divided by the total number of HPV-positive women (*n* = 50,320) × 100%. Abbreviation: HPV, human papillomavirus.

The horizontal stacked bar chart illustrates the overall prevalence (%) across 21 HPV genotypes, ranked in descending order by any-type estimate (Any.). The total length of each horizontal bar represents the overall prevalence burden of a specific genotype in the screened population, further subdivided into single-genotype infections (blue segments) and multiple infections (orange segments). The data indicate that HPV52, HPV58, and HPV16 are the top three predominant genotypes in Ningbo. Moreover, across the vast majority of genotypes—including the leading ones-single-genotype infections account for the predominant proportion of the overall infection burden. HPV: human papillomavirus; Any.: any-type estimate.

To verify whether variations in diagnostic performance across commercial assays introduced bias, sensitivity analyses were performed for the predominant regional oncogenic genotypes (HPV52 and HPV58) within the primary target population requiring precise clinical triage (women with ASC-US/LSIL). Because the BGI Genomics assay yielded no eligible follow-up cases in this specific subcohort, it was excluded from this sensitivity analysis. Among the remaining three mainstream PCR assays, no statistically significant differences were observed in the immediate CIN2 + risk for HPV52-positive women (Dian: 5.37%, MedicalSystem: 5.77%, and Liferiver: 4.28%; $\chi^2 = 1.50$, $P = 0.472$). Similarly, the immediate risk of CIN3 + demonstrated high inter-assay consistency ($P = 0.536$). Consistently, HPV58 also exhibited highly comparable risk estimates across platforms for both CIN2+ ($\chi^2 = 4.24$, $P = 0.120$) and CIN3+ ($P = 0.518$) (Supplementary Table A.3). These findings confirm strong homogeneity across the multiple assays used in this region when evaluating high-grade lesion risks, providing a robust methodological rationale for pooling the overall cohort data in subsequent attribution analyses [10].

3.2. Overall immediate risk of CIN2 + for HPV16/18

Prior to investigating the oncogenic heterogeneity of non-16/18 hrHPV genotypes under ASC-US/LSIL cytological backgrounds, we first evaluated the overall clinical risk of the highest-risk genotypes (HPV16/18) in our cohort to establish a clear clinical benchmark. According to the cervical cancer screening pathway in Ningbo, women testing positive for HPV16/18 were directly referred to colposcopy without cytological triage. Consequently, we calculated their overall immediate risk of CIN2 + regardless of cytological status. As shown in Table 2, HPV16 and HPV18 exhibited potent oncogenic potential in our cohort. Specifically, under the Any. model, the overall immediate risk of CIN2 + reached 13.41% (692/5,159) for HPV16 and 4.05% (120/2,965) for HPV18. Both values exceeded the 4% colposcopy referral threshold recommended by the American Society for Colposcopy and Cervical Pathology (ASCCP) guidelines, reaffirming their clinical status as the highest-risk genotypes. Using this high-risk benchmark, we further quantified the immediate CIN2 + risks of individual extended non-16/18 hrHPV genotypes under ASC-US/LSIL cytological backgrounds to identify specific genotypes with oncogenic potential approaching that of HPV16/18.

Table 2
Overall immediate risk of CIN2 + for HPV16/18 in the cervical cancer screening cohort in Ningbo, China.

HPV genotype	Positive cases(*n*)	confirmed CIN2 + cases(*n*)	Immediate risk of CIN2+(%)
HPV16	5,159	692	13.41
HPV18	2,965	120	4.05

Note: This table presents the overall immediate risk of CIN2 + among HPV16/18-positive women based on the any-type estimate (Any.). According to the cervical cancer screening strategy in Ningbo, HPV16/18-positive women were directly referred to colposcopy without liquid-based cytology (TCT) triage; thus, the data reflect overall oncogenic risks regardless of cytological abnormality grades. Immediate risk of CIN2+ (%) was calculated as the number of confirmed CIN2 + cases divided by the number of positive cases for each genotype × 100%. Abbreviations: HPV, human papillomavirus; TCT, ThinPrep cytologic test; CIN2+, cervical intraepithelial neoplasia grade 2 or worse (defined as the primary clinical endpoint in this study, encompassing CIN2, CIN3, carcinoma in situ, and invasive cervical cancer).

3.3. Immediate risk of CIN2 + for non-16/18 genotypes under ASC-US and LSIL backgrounds

Although the aforementioned epidemiological data indicate that non-16/18 hrHPV genotypes, such as HPV52 and HPV58, exhibit high prevalence and constitute a large proportion of infections in Ningbo, a high infection burden does not necessarily equate to a high oncogenic risk. In current clinical practice, a uniform triage standard is often applied to all individuals infected with non-16/18 hrHPV, overlooking the pathogenic heterogeneity among individual extended genotypes. To further elucidate the true clinical pathogenicity of these non-16/18 hrHPV genotypes, we specifically evaluated the immediate risks of CIN2 + attributable to different genotypes under ASC-US and LSIL cytological backgrounds. The derived genotype-specific immediate risks were then compared against the ASCCP—recommended 4% colposcopy referral threshold to serve as an evidence base for establishing risk-stratified triage strategies.

To precisely isolate the independent pathogenicity of non-16/18 hrHPV genotypes causing CIN2 + in the ASC-US and LSIL populations, this study employed three statistical approaches for lesion attribution: (1) minimum estimate (Min.), (2) any-type estimate (Any.), and (3) hierarchical attribution estimate (Hier.), combined with odds ratios (ORs) to evaluate relative risk intensity (Supplementary Table A.4 and A.5). In establishing the global priority sequence for the Hier. model, we comprehensively considered the carcinogenic biological classifications by the International Agency for Research on Cancer (IARC) and the actual pathogenic risks observed in our large-scale regional cohort [11]. For sporadic genotypes with excessively small sample sizes ($N < 500$) and a lack of statistical stability (e.g., HPV53, HPV82, and HPV73), even if their immediate risks were spuriously elevated due to statistical bias, we consistently assigned them to a lower priority tier based on their biological attributes to prevent misattribution in cases of multiple infections. Consequently, the specific sequence was defined as follows: HPV31 > 33 > 58 > 52 > 35 > 66 > 59 > 45 > 51 > 56 > 68 > 39 > 82 > 53 > 73 (Supplementary Table A.6). For multiple infections, lesion outcomes were exclusively attributed to the genotype with the highest priority within the patient's concurrent infection profile.

3.3.1. Immediate risk and risk assessment of CIN2 + for non-16/18 genotypes in women with ASC-US

As shown in Table 3 and Fig. 5, the oncogenic risks among individual non-16/18 genotypes exhibited highly significant heterogeneity within the ASC-US population. Notably, HPV31, 33, and 58 constituted the core high-risk tier for driving CIN2 + lesions. Using the average pathogenic risk of the overall non-16/18-positive cohort as a baseline reference, patients infected with HPV31 (OR = 2.04, 95% CI: 1.29–3.24, $P = 0.002$) and HPV33 (OR = 2.09, 95% CI: 1.43–3.05, $P < 0.001$) experienced a significantly doubled immediate risk of CIN2+. HPV58 (OR = 1.54, 95% CI: 1.13–2.10, $P = 0.006$) also demonstrated a markedly elevated relative risk. Despite being the most prevalent genotype in the Chinese population, HPV52 exhibited a pathogenic risk highly consistent with the overall non-16/18 average, showing no statistically significant difference (OR = 1.09, 95% CI: 0.82–1.44, $P = 0.846$). This finding aligns with previous studies conducted in Ningbo and West China Hospital, Sichuan [12–13]. Conversely, genotypes at the lower end of the spectrum displayed pronounced low-risk characteristics; in particular, the relative risks for HPV39 (OR = 0.33, 95% CI: 0.12–0.89, $P = 0.021$) and HPV51 (OR = 0.56, 95% CI: 0.32–0.97, $P = 0.035$) were significantly lower than the baseline. This suggests that these specific genotypes possess limited capacity to drive high-grade lesion progression in an ASC-US background, thereby lacking statistical justification for prioritized colposcopy referral.

Table 3

Immediate risk and risk assessment of CIN2 + for non-16/18 HPV genotypes in the ASC-US population.

HPV genotype	Immediate risk of CIN2 + [a] (%)			Relative risk assessment[b]	
	Min.	Any.	Hier.	OR(95% CI)	*P value*
HPV31	10.85	9.65	9.65	2.04(1.29, 3.24)	0.002
HPV33	10.50	9.54	9.50	2.09(1.43, 3.05)	< 0.001
HPV58	6.50	7.14	6.97	1.54(1.13, 2.10)	0.006
HPV59	5.60	4.57	5.19	0.86(0.45, 1.65)	0.655
HPV52	4.84	5.48	4.85	1.09(0.82, 1.44)	0.846
HPV35	4.88	5.50	4.38	1.06(0.58, 1.93)	0.846
HPV56	3.53	4.34	3.63	0.81(0.47, 1.38)	0.435
HPV45	4.23	4.00	3.53	0.75(0.30, 1.85)	0.531
HPV68	2.36	3.10	2.18	0.56(0.31, 1.01)	0.052
HPV51	2.08	3.14	1.79	0.56(0.32, 0.97)	0.035
HPV66	1.85	2.75	1.33	0.51(0.16, 1.61)	0.229
HPV39	0.00	1.85	0.00	0.33(0.12, 0.89)	0.021
HPV53	0.00	0.00	0.00	NA[c]	-
HPV73	0.00	0.00	0.00	NA	-
HPV82	0.00	0.00	0.00	NA	-

Note: HPV genotypes in the table are ranked in descending order based on their immediate risk of CIN2 + under the Hier. model.

[a] Immediate risk assessment: Calculated using three epidemiological algorithms: Min. (minimum estimate, including only women with single-genotype HPV infections); Any. (any-type estimate, including all instances of HPV infections, where multiple infections are counted repeatedly); and Hier. (hierarchical attribution estimate, exclusively attributing multiple infections to the specific genotype with the highest risk ranking within the Any. model).

[b] Relative risk assessment: Relative risks for individual genotypes were calculated using the average risk of the overall non-16/18-positive cohort as the baseline reference (OR = 1.0). Statistical inferences in this study were based on the Any. algorithm.

[c] Data notes: For genotypes with no detected CIN2 + cases in this cohort (e.g., HPV53 and HPV82), the immediate risk across all algorithms was recorded as 0.00%; their ORs are denoted as NA because valid statistical regression could not be performed. *P* values were derived from Pearson's chi-square test without continuity correction by default.

Abbreviations: HPV, human papillomavirus; CIN2+, cervical intraepithelial neoplasia grade 2 or worse (defined as the primary positive endpoint in this study); OR, odds ratio; CI, confidence interval; NA, not applicable/not available.

The figure illustrates the confounding effect of multiple infections on the risk assessment of individual genotypes. (A) Forest plot displaying the relative risks (ORs and 95% CIs) of CIN2 + driven by individual genotypes, using the average risk of the overall non-16/18-positive cohort as the baseline reference (OR = 1.0). The vertical solid line represents the null line (OR = 1.0). (B) Dumbbell plot showing the immediate CIN2 + risks (%) for each genotype in the ASC-US cohort, calculated via three statistical approaches. Blue circles, green squares, and orange triangles denote the immediate CIN2 + risks (%) under the Any., Hier., and Min. algorithms, respectively. The horizontal connecting line for each genotype represents the risk assessment span across different methods, visually reflecting the degree to which multiple infections interfere with the evaluation of a genotype's true independent risk. A wider span indicates a more pronounced confounding effect from multiple infections. Abbreviations: HPV, human papillomavirus; OR, odds ratio; CI, confidence interval.

3.3.2. Immediate risk and risk assessment of CIN2 + for non-16/18 genotypes in women with LSIL

As the severity of cytological abnormalities progressed from ASC-US to LSIL, the oncogenic drive of non-16/18 hrHPV genotypes shifted accordingly. Notably, the core pathogenic genotypes demonstrated a pronounced "risk aggregation effect" within this higher-grade cytological context (Table 4 and Fig. 6). Compared to the ASC-US cohort, the relative risks of HPV31 and HPV58 surged in the LSIL background, with their ORs rising to 2.81 (95% CI: 1.43–5.54, $P = 0.002$) and 2.43 (95% CI: 1.58–3.72, $P < 0.001$), respectively. This indicates that once mild morphological changes (LSIL) manifest, the independent oncogenic capacity of these two genotypes to drive progression toward high-grade lesions becomes exceptionally potent. Concurrently, the "risk marginalization" of the lower-tier genotypes was further amplified in the LSIL background. Most representatively, the ORs for HPV56 and HPV68 plummeted to 0.25 (95% CI: 0.09–0.68, $P = 0.004$) and 0.21 (95% CI: 0.05–0.85, $P = 0.016$), respectively. This demonstrates that even when patients infected with these specific genotypes develop LSIL, their risk of progressing to CIN2 + remains significantly diminished. In the LSIL background, HPV82 exhibited an exceptionally high immediate risk of CIN2+ (20.00% under the Any. algorithm). However, this was primarily attributable to the extremely small overall infection base of HPV82 in the regional cohort (only 10 cases), which caused isolated cases of high-grade lesions to disproportionately amplify the overall risk proportion. This observation aligns with the findings of Li Xiao et al., who also reported a high risk for HPV82 (50.00%) that was similarly limited by statistical bias stemming from a small sample size [14].

Table 4
Immediate risk and risk assessment of CIN2 + for non-16/18 HPV genotypes in the LSIL population.

HPV genotype	Immediate risk of CIN2 + [a](%)			Relative risk assessment[b]	
	Min.	Any.	Hier.	OR(95% CI)	*P value*
HPV31	15.91	15.94	15.94	2.81(1.43, 5.54)	0.002
HPV58	15.43	12.37	11.85	2.43(1.58, 3.72)	< 0.001
HPV33	9.57	8.11	8.16	1.25(0.67, 2.34)	0.485
HPV52	6.97	7.26	6.49	1.13(0.75, 1.70)	0.544
HPV51	6.87	5.49	5.56	0.78(0.43, 1.41)	0.403
HPV39	5.26	3.70	5.26	0.52(0.16, 1.67)	0.263
HPV35	4.65	4.55	5.17	0.65(0.23, 1.80)	0.398
HPV66	2.08	3.45	3.23	0.48(0.15, 1.54)	0.208
HPV59	0.00	3.39	2.44	0.48(0.11, 1.97)	0.296
HPV56	1.63	1.97	1.48	0.25(0.09, 0.68)	0.004
HPV45	0.00	0.00	0.00	NA[c]	-
HPV53	0.00	0.00	0.00	NA	-
HPV68	0.00	1.59	0.00	0.21(0.05, 0.85)	0.016
HPV73	0.00	0.00	0.00	NA	-
HPV82	0.00	20.00	0.00	3.51(0.73.16.71)	0.093

Note: HPV genotypes in the table are ranked in descending order based on their immediate risk of CIN2 + under the Hier. model.

[a] Immediate risk assessment: Calculated using three epidemiological algorithms: Min. (minimum estimate, including only women with single-genotype HPV infections); Any. (any-type estimate, including all instances of HPV infections, where multiple infections are counted repeatedly); and Hier. (hierarchical attribution estimate, exclusively attributing multiple infections to the specific genotype with the highest risk ranking within the Any. model).

[b] Relative risk assessment: Relative risks for individual genotypes were calculated using the average risk of the overall non-16/18-positive cohort as the baseline reference (OR = 1.0). Statistical inferences in this study were based on the Any. algorithm.

[c] Data notes: For genotypes with no detected CIN2 + cases in this cohort (e.g., HPV45 and HPV53), the immediate risk across all algorithms was recorded as 0.00%; their ORs are denoted as NA because valid statistical regression could not be performed. *P* values were derived from Pearson's chi-square test without continuity correction by default.

Abbreviations: HPV, human papillomavirus; CIN2+, cervical intraepithelial neoplasia grade 2 or worse (defined as the primary positive endpoint in this study); OR, odds ratio; CI, confidence interval; NA, not applicable/not available.

The figure illustrates the confounding effect of multiple infections on the risk assessment of individual genotypes. (A) Forest plot displaying the relative risks (ORs and 95% CIs) of CIN2 + driven by individual genotypes, using the average risk of the overall non-16/18-positive cohort as the baseline reference (OR = 1.0). The vertical solid line represents the null line (OR = 1.0). (B) Dumbbell plot showing the immediate CIN2 + risks (%) for each genotype in the LSIL cohort, calculated via three statistical approaches. Blue circles, green squares, and orange triangles denote the immediate CIN2 + risks (%) under the Any., Hier., and Min. algorithms, respectively. The horizontal connecting line for each genotype represents the risk assessment span across different methods, visually reflecting the degree to which multiple infections interfere with the evaluation of a genotype's true independent risk. A wider span indicates a more pronounced confounding effect from multiple infections. Abbreviations: HPV, human papillomavirus; OR, odds ratio; CI, confidence interval.

Taken together, these comparative analyses demonstrate that as cytological abnormalities progress from ASC-US to LSIL, the oncogenic capacity across different HPV genotypes exhibits a progressive evolution of risk heterogeneity. This dynamic shift provides the most direct evidence for implementing differentiated, risk-stratified clinical triage strategies.

3.4. Risk-stratified triage strategy for non-16/18 genotypes under ASC-US/LSIL cytology

To formulate a more precise management protocol for women with ASC-US or LSIL cytology, and guided by the ASCCP "equal risk, equal management" consensus [15], we propose a three-tier triage strategy for non-16/18 hrHPV genotypes:

- • **Extremely high-risk tier (HPV31, 33, 52, 58)**: Immediate risks consistently exceeded the 4% ASCCP threshold regardless of an ASC-US or LSIL background. We recommend direct colposcopy referral to prevent missed diagnoses. Although HPV52 presented a moderate relative risk, its absolute risk remained robustly above the threshold owing to a massive infection base, confirming the potent oncogenic drive of this entire tier.
- • **Intermediate-high-risk tier (HPV35, 45, 51, 59, 82)**: This group exhibited substantial risk heterogeneity and statistical uncertainty (e.g., radical fluctuations in HPV82 driven by small-sample bias). For these women, incorporating second-line triage methods (such as DNA methylation or p16/Ki-67 dual staining) is advised to accurately identify individuals at genuine risk of high-grade lesions.
- • **Low-risk tier (HPV39, 56, 66, 68)**: Immediate risks predominantly fell below 4%. Notably, HPV39 risk approached zero in ASC-US cases, and HPV56/68 showed significantly diminished relative risks in LSIL cases. For this tier, we recommend deferring colposcopy and performing a 1-year follow-up instead (Table 5).

Table 5
Refined management recommendations for non-16/18 HPV-positive patients based on extended genotyping.

Risk stratification tier	Included HPV genotypes	Immediate risk range of CIN2+ (%)	Immediate risk characteristics of CIN2+ (ASC-US/LSIL)	Recommended clinical management strategy
Prioritized referral group (Extremely high-risk)	HPV31, 33, 52, 58	4.85–15.94	Consistently > 4%	Direct referral to colposcopy
Adjunctive triage group (Intermediate-high-risk)	HPV35, 45, 51, 59, 82	0.00-5.56	High risk heterogeneity and high statistical uncertainty	Introduce second-line molecular biomarkers
Conservative follow-up groups (Low-risk)	HPV39, 56, 66, 68	0.00-5.26	Predominantly and consistently < 4%	Defer colposcopy, 1-year follow-up

Note: HPV, human papillomavirus; CIN2+, cervical intraepithelial neoplasia grade 2 or worse (defined as the primary positive endpoint in this study).

The traditional strategy of referring all non-16/18 hrHPV-positive ASC-US/LSIL women for colposcopy achieved high sensitivity but zero specificity (0.00%), resulting in a massive misallocation of medical resources (5,692 total referrals). In contrast, our updated risk-based strategy—referring only patients infected with the extremely high-risk genotypes (HPV31, 33, 52, and 58)—kept missed diagnoses well within a safe margin, yielding an NPV of 96.77%. The new protocol not only exhibited significantly higher specificity (40.11% vs. 0.00%, $P < 0.001$) and PPV (7.21% vs. 5.66%, $P < 0.001$) (Table 6) but also safely bypassed 39.1% (2,226/5,692) of unnecessary colposcopic exams. Furthermore, it provided substantial health-economic benefits by reducing the number of colposcopies needed to detect one CIN2 + lesion from 17.68 to 13.86. Ultimately, integrating extended HPV genotyping into a risk-stratified triage framework successfully establishes the optimal balance between accurate high-grade lesion detection and the minimization of overtreatment.

Table 6
Comparison of two triage strategies for detecting CIN2 + lesions in non-16/18 hrHPV-positive women with ASC-US/LSIL.

Evaluation metric	Traditional strategy (Universal referral)	New strategy (Risk-stratified)	*P* value
Colposcopy referrals (*n*)	5,692	3,466	< 0.001
Sensitivity (%)	100.00	77.64	< 0.001
Specificity (%)	0.00	40.11	< 0.001
Positive predictive value (PPV, %)	5.66	7.21	0.003
Negative predictive value (NPV, %)	--	96.77	--
Number of colposcopies needed to detect one CIN2 + cases (NNC)	17.68	13.86	0.003

Note: Traditional strategy: Immediate colposcopy referral for all non-16/18 hrHPV-positive women with ASC-US/LSIL. New strategy: Immediate colposcopy referral only for ASC-US/LSIL women infected with extremely high-risk genotypes (HPV31, 33, 52, and 58). *P* values were calculated using McNemar's test for sensitivity and specificity, and the two-sample test for proportions for PPV and NNC. Abbreviations: HPV, human papillomavirus; hrHPV: high-risk human papillomavirus; CIN2+, cervical intraepithelial neoplasia grade 2 or worse.

4. Discussion

Leveraging a large-scale real-world dataset of over 510,000 cases, this study demonstrated substantial epidemiological and pathogenic heterogeneity among non-16/18 hrHPV infections within the ASC-US/LSIL population. In the Ningbo region, single-genotype infections predominated, with HPV52, 58, and 16 emerging as the top three dominant genotypes. This aligns with the epidemiological profiles observed in other southern Chinese cities (such as Suzhou and Guangzhou) as well as in Singapore [1, 16–18]. However, the high local prevalence of HPV52 and 58 highlights a distinct regional disparity compared to Western countries, where HPV31 and 51 are typically more prevalent [19].

Regarding age distribution, the HPV infection pattern in our cohort deviated from the classic "U-shaped" curve [16]. Constrained by the age threshold of the national screening program (which primarily targets women aged 35–64 years), the overall HPV infection rate exhibited a gradual upward trend with advancing age. The peak infection rate emerged in the postmenopausal demographic (≥ 60 years, 12.79%), contrasting with the peak at "around 50 years of age" reported in other literature [18]. This discrepancy may be associated with the regional socioeconomic status of the study population [20–21], or potentially driven by the reactivation of latent HPV infections secondary to declining immune function after menopause [22]. Crucially, unlike hospital-based cohorts comprising symptomatic outpatients, our study is grounded in an asymptomatic, community-based screening population, thereby more accurately reflecting the true epidemiological characteristics of the general public. Given these population-specific risk profiles, our findings strongly caution against blindly adopting management guidelines derived from Western epidemiological data or triage experiences based on outpatient cohorts. Instead, it is imperative to construct refined, region-specific triage strategies grounded in localized real-world data. Such a community-

derived approach effectively circumvents the inherent selection bias of outpatient cohorts, demonstrating superior generalizability and health-economic efficiency for guiding population-scale screening programs [23].

This study profoundly elucidates the interplay between cytological severity and HPV pathogenicity. As the cytological background progresses from ASC-US to LSIL, the internal risk gradient among non-16/18 hrHPV genotypes undergoes significant divergence: the relative risks for highly pathogenic types such as HPV31 and 58 escalate further, whereas the relative risks for lower-tier types like HPV56 concurrently decline. Although ASC-US and LSIL carry distinct clinical implications—with some guidelines conventionally recommending direct referral or shortened follow-up intervals for LSIL—this study nevertheless proposes adopting a unified, extended genotyping-based triage framework for both cytological contexts, specifically designating HPV31, 33, 52, and 58 as the extremely high-risk tier. This classification is driven not only by our statistical findings, where the immediate risks of all four genotypes universally breached the 4% clinical threshold, but also by a solid biological foundation. Phylogenetically, these four genotypes belong to the Alpha-9 species, sharing high sequence homology with HPV16, which intrinsically endows them with a potent oncogenic drive [24]. In stark contrast to large-scale Western cohorts (such as the ATHENA trial, which reported lower risks for HPV52 and 58), our study further corroborates that HPV52 and 58 exhibit exceptionally high pathogenicity comparable to that of HPV31 and 33 in the Chinese population [25]. Consequently, in domestic clinical practice, patients infected with these four extremely high-risk genotypes warrant immediate colposcopic evaluation regardless of an ASC-US or LSIL background, rather than being subjected to delayed intervention as suggested by conventional Western guidelines.

Building upon the aforementioned risk assessments, anchored by the ASCCP consensus threshold (immediate risk ≥ 4% for colposcopy referral vs. <4% for routine follow-up), and actively responding to the latest WHO guidelines advocating for tailored clinical pathways based on extended HPV genotyping [6, 14, 26], this study proposes a refined, three-tier risk-stratified management strategy. This framework introduces several critical distinctions from current mainstream international guidelines and Western-centric practices. First, regarding the extremely high-risk tier (HPV31, 33, 52, and 58), unlike certain Western studies that heavily emphasize the risk of HPV45 [27], our regional data demonstrate that the oncogenic drive of HPV58 is exceptionally prominent. This discrepancy is likely attributable to the circulation of specific, highly pathogenic HPV58 sub-lineages uniquely prevalent in Asian populations [28]. Consequently, we strongly advocate that patients in this tier be managed analogously to those with HPV16/18 infections—prompting direct colposcopy referral to control the risk of missed diagnoses [27]. Second, for the intermediate-high-risk tier (HPV35, 45, 51, 59, and 82), a singular cytological background can no longer provide a reliable basis for risk stratification, as evidenced by the high statistical heterogeneity and small-sample fluctuations. Diverging from conventional triage reliant solely on cytology or viral load, we recommend integrating second-line molecular biomarkers (such as DNA methylation, p16/Ki-67 dual staining, or specific HPV E6/E7 mRNA assays) for precise secondary triage [29–31]. This aligns robustly with the prevailing trends of the international NTCC2 trial [32]. Finally, for the low-risk tier (HPV39, 56, 66, and 68), whose independent pathogenic risks consistently remained at low levels, our strategy breaks away from the traditional “one-size-fits-all” paradigm, recommending conservative management with a 1-year routine follow-up. The clinical and health-economic superiority of this differentiated pathway is directly reflected in our

performance metrics. Although the traditional uniform referral strategy achieved the highest sensitivity, its absolute lack of specificity (0.00%) resulted in a massive misallocation of medical resources. Our proposed protocol proves that by deferring immediate colposcopy for relatively low-risk subtypes (e.g., HPV39, 56, 66, and 68), specificity can be drastically improved by 40.11%. Furthermore, the significant optimization of the number of colposcopies needed to detect one CIN2 + case (NNC dropping from 17.68 to 13.86) substantially elevates the health-economic efficiency of the screening program, successfully maximizing the mitigation of overtreatment while ensuring clinical safety.

The primary strength of this study lies not only in its massive sample size but also in its authentic reflection of the diverse testing modalities utilized in routine clinical practice. Rather than relying on idealized laboratory data derived from a single proprietary assay, our conclusions encapsulate the pragmatic real-world performance of various approved commercial kits deployed concurrently across a vast population. Consequently, the triage strategy derived from these data boasts robust generalizability and superior translational value for broader clinical implementation [33]. Methodologically, we rigorously minimized the confounding bias inherently introduced by multiple infections by incorporating three distinct epidemiological approaches: the minimum, any, and hierarchical attribution estimates [8].

Nonetheless, several limitations of this study must be acknowledged. First, although our preceding sensitivity analyses confirmed the consistency of risk assessments for the primary pathogenic genotypes, the concurrent use of diverse PCR diagnostic kits inevitably introduces certain biases stemming from variations in detection thresholds and primer designs. Second, inherently restricted by its cross-sectional design, this study primarily evaluated the immediate risk—which exclusively reflects the concurrent lesion burden at the time of initial screening—thereby precluding the dynamic tracking of clearance or progression trajectories for non-16/18 hrHPV infections [34]. As evidenced by a prospective cohort study in the United States, even among women initially negative for high-grade lesions, the long-term persistent infection of specific high-risk genotypes (e.g., HPV16, 18, 31, and 33) still triggers a drastic escalation in subsequent risk [35]. This phenomenon warrants particular vigilance regarding the massive HPV52-positive subpopulation. In our cohort, HPV52 accounted for the absolute majority of infections (33.65%). While its relative pathogenicity hovered around the baseline average, its absolute risk firmly breached the 4% clinical intervention threshold. Furthermore, although HPV52 infections are predominantly transient, previous literature indicates that its oncogenic process is notably more indolent than that of HPV16. Specifically, a subset of HPV52 infections accompanied by low-grade morphological changes may require years of persistence before progressing to CIN2 + or worse [13]. This dictates that the cumulative risk of HPV52 will ultimately surpass its immediate risk, implying that our current cross-sectional data might underestimate its long-term pathogenic threat. Consequently, while the primary triage strategy proposed herein effectively addresses the immediate dilemma of over-referral within a massive positive population, future prospective studies are urgently needed. Such studies should explore the feasibility of integrating molecular biomarkers to dynamically monitor highly prevalent yet slowly progressing genotypes like HPV52. Finally, constrained by the accessibility of real-world retrospective data, our study was unable to incorporate potential confounding variables such as patients' HPV vaccination history, sexual behaviors, and smoking status. Because these covariates may significantly influence the subsequent progression of specific HPV genotypes, further prospective validation is required to definitively isolate their independent oncogenic effects.

5. Conclusions

In a massive real-world cohort of 510,000 screened women in the Ningbo region, although non-16/18 hrHPV infections (particularly HPV52 and 58) predominantly drove ASC-US and LSIL cytology, the immediate risks of CIN2 + exhibited profound genotype-specific heterogeneity. Our findings underscore that non-16/18 genotypes must not be managed as a homogeneous entity. The extremely high-risk tier (HPV31, 33, 52, and 58) demonstrated a potent oncogenic drive; therefore, we strongly advocate elevating their triage priority to parallel that of HPV16 and 18, warranting direct colposcopy referral. Conversely, for ASC-US/LSIL populations infected with low-risk genotypes (HPV39, 56, 66, and 68) yielding immediate risks below the 4% threshold, deploying a conservative follow-up strategy is both clinically safe and health-economically sound. Ultimately, implementing this three-tier, risk-stratified triage strategy based on extended HPV genotyping successfully strikes the optimal balance between maximizing the detection of high-grade lesions and mitigating overtreatment. This framework holds substantial translational value for primary screening triage, providing robust evidence-based support for the precision management of cervical cancer.

Abbreviations

CIN
cervical intraepithelial neoplasia
hrHPV
high-risk human papillomavirus
ASC-US
atypical squamous cells of undetermined significance
LSIL
low-grade squamous intraepithelial lesion
WHO
the World Health Organization
ECC
endocervical curettage
lrHPV
low-risk human papillomavirus
AIS
adenocarcinoma in situ
VaIN
vaginal intraepithelial neoplasia
Any.
any-type estimate
Min.
minimum estimate
Hier.
Hierarchical-attribution estimate
OR

odds ratio
PPV
positive predictive value
NPV
negative predictive value
NNC
the number of colposcopies needed to detect one CIN2 + case (NNC)
IARC
the International Agency for Research on Cancer

Declarations

Ethics approval and consent to participate

This study protocol was approved by the Medical Ethics Committee of Ningbo University Affiliated Women and Children's Hospital (Approval No.: NBFE-2026-KY-005). Given the retrospective nature of this study, which utilized de-identified historical screening records, the requirement for obtaining written informed consent from individual patients was waived by the Medical Ethics Committee of Ningbo University Affiliated Women and Children's Hospital.

Consent for publication

Not applicable.

Funding

This work was supported by the Ningbo Top Medical and Health Research Program [grant numbers: No. 2024021020]; the Ningbo Clinical Medical Research Center for Gynecological Diseases [grant numbers: No. 2024L002]; and the Zhejiang Provincial Key Clinical Specialty of Gynecology.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author used Gemini to translate and polish the English language. After using this tool, the author reviewed and edited the content as needed and take full responsibility for the content of the published article.

Author Contribution

Conceptualization, J.Z. and X.L.; methodology, Y.C. and Ju.Z.; data curation, L.C. and Li.C.; formal analysis, J.M. and L.W.; writing-original draft preparation, J.M. and L.W.; writing-review and editing, J.Z. and X.L.; supervision, J.Z. and X.L. All authors have read and agreed to the published version of the manuscript.

Acknowledgement

The authors thank all medical staff for clinical data collection, the patients for their participation, and the research team for valuable contributions to this study and manuscript preparation.

Data Availability Statement

The datasets generated and analysed in this study are not publicly available owing to patient confidentiality and ethical constraints, but can be obtained from the corresponding authors (Ma Jun and Zhang Jing) upon reasonable request and with permission of the Medical Ethics Committee of Ningbo University Affiliated Women and Children's Hospital.

Competing interests

No potential conflict of interest was reported by the authors.

Acknowledgements

The authors thank all medical staff for clinical data collection, the patients for their participation, and the research team for valuable contributions to this study and manuscript preparation.

References

1. WANG T, LUAN L, DENG J, et al. Prevalence and human papillomavirus (HPV) genotype distribution in Suzhou, China [J]. Hum Vaccin Immunother. 2023;19(2):2241309.
2. MASSAD L S, CLARKE M A, PERKINS R B, et al. Applying Results of Extended Genotyping to Management of Positive Cervicovaginal Human Papillomavirus Test Results: Enduring Guidelines [J]. J Low Genit Tract Dis. 2025;29(2):134–43.
3. ZHANG Y, LI H, YOU Q, et al. Decoding Fujian's cervical HPV landscape: unmasking dominance of non-16/18 HR-HPV and tailoring prevention strategies at a large scale [J]. Front Public Health. 2024;12:1357073.
4. SU H, JIN X. The triage role of cytological DNA methylation in women with non-16/18, specifically genotyping high-risk HPV infection [J]. Br J Cancer. 2025;132(11):1064–71.
5. WANG Z, GU Y, WANG H, et al. Distribution of cervical lesions in high-risk HPV (hr-HPV) positive women with ASC-US: a retrospective single-center study in China [J]. Virol J. 2020;17(1):185.
6. WHO Guidelines Approved by the Guidelines Review Committee [M]. WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention: Use of human papillomavirus (HPV) DNA genotyping. Geneva; World Health Organization © World Health Organization 2026. 2026.
7. CREE I A, WHITE V A, INDAVE B I, et al. Revising the WHO classification: female genital tract tumours [J]. Histopathology. 2020;76(1):151–6.

8. YU Y, HAO J, MOHAMED S B, et al. The prevalence of multiple or single HPV infection and genotype distribution in healthy Chinese women: A systemic review [J]. J Cancer Res Ther. 2024;20(4):1265–73.
9. LI X, RAO X, WEI M J et al. Extended HPV Genotyping for Risk Assessment of Cervical Intraepithelial Neoplasia Grade 2/3 or Worse in a Cohort Study [J]. J Natl Compr Canc Netw, 2022, 20(8): 906 – 14.e10.
10. VOLESKY K D MAGNANS, MAYRAND M H, et al. Clinical Performance of the BD Onclarity Extended Genotyping Assay for the Management of Women Positive for Human Papillomavirus in Cervical Cancer Screening [J]. Cancer Epidemiol Biomarkers Prev. 2022;31(4):851–7.
11. STOLER M H, PARVU V, YANSON K, et al. Risk stratification of HPV-positive results using extended genotyping and cytology: Data from the baseline phase of the Onclarity trial [J]. Gynecol Oncol. 2023;174:68–75.
12. CHEN S, HU S, YIN J, et al. Ranking the attribution of high-risk genotypes among women with cervical precancers and cancers: a cross-sectional study in Ningbo, China [J]. Infect Agent Cancer. 2024;19(1):43.
13. JIANG W, AUSTIN R M, ZHANG H, et al. The Clinical Utility of Extended High-Risk HPV Genotyping in Women With ASC-US Cytology [J]. Am J Clin Pathol. 2022;158(4):472–9.
14. YE C, LIU Y, HUANG H, et al. Risk-Based Triage Strategy by Extended HPV Genotyping for Women With LSIL Cytology: A Real-World Study [J]. J Med Virol. 2025;97(5):e70404.
15. PERKINS R B, GUIDO R S, CASTLE PE, et al. 2019 ASCCP Risk-Based Management Consensus Guidelines for Abnormal Cervical Cancer Screening Tests and Cancer Precursors [J]. J Low Genit Tract Dis. 2020;24(2):102–31.
16. HAN S, LIN M, LIU M, et al. Prevalence, trends, and geographic distribution of human papillomavirus infection in Chinese women: a summative analysis of 2,728,321 cases [J]. BMC Med. 2025;23(1):158.
17. NANDAKUMAR V, CHANG C P KRISHNANS, et al. High-risk human papillomavirus genotype prevalence in Singapore cervical samples: Implications for clinical management [J]. Ann Acad Med Singap. 2025;54(12):758–65.
18. YANG X, LI Y, TANG Y, et al. Cervical HPV infection in Guangzhou, China: an epidemiological study of 198,111 women from 2015 to 2021 [J]. Emerg Microbes Infect. 2023;12(1):e2176009.
19. TSAKOGIANNIS D, ZOGRAFOS E. TZIOGA L, Prevalence and Genotype Distribution of High-Risk HPV Genotypes Among Women in Greece: A Retrospective Analysis of 3500 Women [J]. Cancers (Basel), 2025, 17(8).
20. LU Z, DONG B, YAO P, et al. Characteristics of Human Papillomavirus Prevalence and Infection Patterns Among Women Aged 35–65 in Fujian Province, China: A Nine-Year Retrospective Observational Study [J]. J Med Virol. 2025;97(2):e70238.
21. YAN M, YANG Y, ZHAO Y, et al. Prevalence and genotype distribution of high-risk human papillomavirus infection among Chinese women aged 35–64: A national screening population-based study [J]. J Infect. 2026;92(3):106697.
22. OSMANI V, ROSSITER M. HÖRNER L, Worldwide burden of cervical human papillomavirus (HPV) in women over 50 years with abnormal cytology: a systematic review and meta-analysis [J]. BMJ Glob

Health, 2025, 10(4).
23. SMITH L W GOTTSCHLICHA. HPV Extended Genotyping to Triage Abnormal Cervical Cancer Screens-Balancing the Harms and Benefits of an Additional Triage Test before Direct Colposcopy Referral [J]. Cancer Epidemiol Biomarkers Prev. 2024;33(8):979–81.
24. YUAN H, YAN Z, GAN J, et al. Phylogenetic analysis and antigenic epitope prediction for E6 and E7 of Alpha-papillomavirus 9 in Taizhou, China [J]. BMC Genomics. 2024;25(1):507.
25. YE Y, JONES T, WANG T et al. Comprehensive overview of genotype distribution and prevalence of human papillomavirus in cervical lesions [J]. Gynecol Obstet Clin Med, 2024, 4(1).
26. RAO X, WANG Y H, CHEN R Z, et al. Risk-based triage strategy by extended HPV genotyping for women with ASC-US cytology [J]. Ann Med. 2025;57(1):2451183.
27. WEI F, GEORGES D, MAN I, et al. Causal attribution of human papillomavirus genotypes to invasive cervical cancer worldwide: a systematic analysis of the global literature [J]. Lancet. 2024;404(10451):435–44.
28. CHEN Z, HO W C S, BOON SS et al. Ancient Evolution and Dispersion of Human Papillomavirus 58 Variants [J]. J Virol, 2017, 91(21).
29. HARPER D M, ANDERSON R J BAKERE, et al. Cost-effectiveness of p16/Ki-67 Dual-Stained Cytology Reflex Following Co-testing with hrHPV Genotyping for Cervical Cancer Screening [J]. Cancer Prev Res (Phila). 2023;16(7):393–404.
30. SALTA S, LOBO J, MAGALHÃES B, et al. DNA methylation as a triage marker for colposcopy referral in HPV-based cervical cancer screening: a systematic review and meta-analysis [J]. Clin Epigenetics. 2023;15(1):125.
31. AL-SHIBLI K, NGUYEN D T, MOHAMMED H A L et al. Genotype-Specific HPV E6/E7 mRNA Triage Improves Risk Stratification and Reduces Referrals in DNA-Positive ASC-US/LSIL: A Real-World Cohort from Nordland. Nor [J] Pathogens, 2026, 15(2).
32. BENEVOLO M, RONCO G, MANCUSO P, et al. Comparison of HPV-positive triage strategies combining extended genotyping with cytology or p16/ki67 dual staining in the Italian NTCC2 study [J]. EBioMedicine. 2024;104:105149.
33. DE MARCO L, BISANZI S, RONCO G, et al. Extended HPV genotyping by the BD Onclarity assay: concordance with screening HPV-DNA assays, triage biomarkers, and histopathology in women from the NTCC2 study [J]. Microbiol Spectr. 2025;13(1):e0089724.
34. ÖZEN S, ÖZEN E G, ŞIMŞEK G, et al. HPV persistence and CIN2 + Detection in women with normal cytology and Non-16/18 High-Risk HPV positivity [J]. BMC Womens Health. 2026;26(1):98.
35. BUKOWSKI A, HOYO C, HUDGENS M G, et al. Extended Human Papillomavirus Genotyping to Predict Progression to High-Grade Cervical Precancer: A Prospective Cohort Study in the Southeastern United States [J]. Cancer Epidemiol Biomarkers Prev. 2022;31(8):1564–71.

Figures

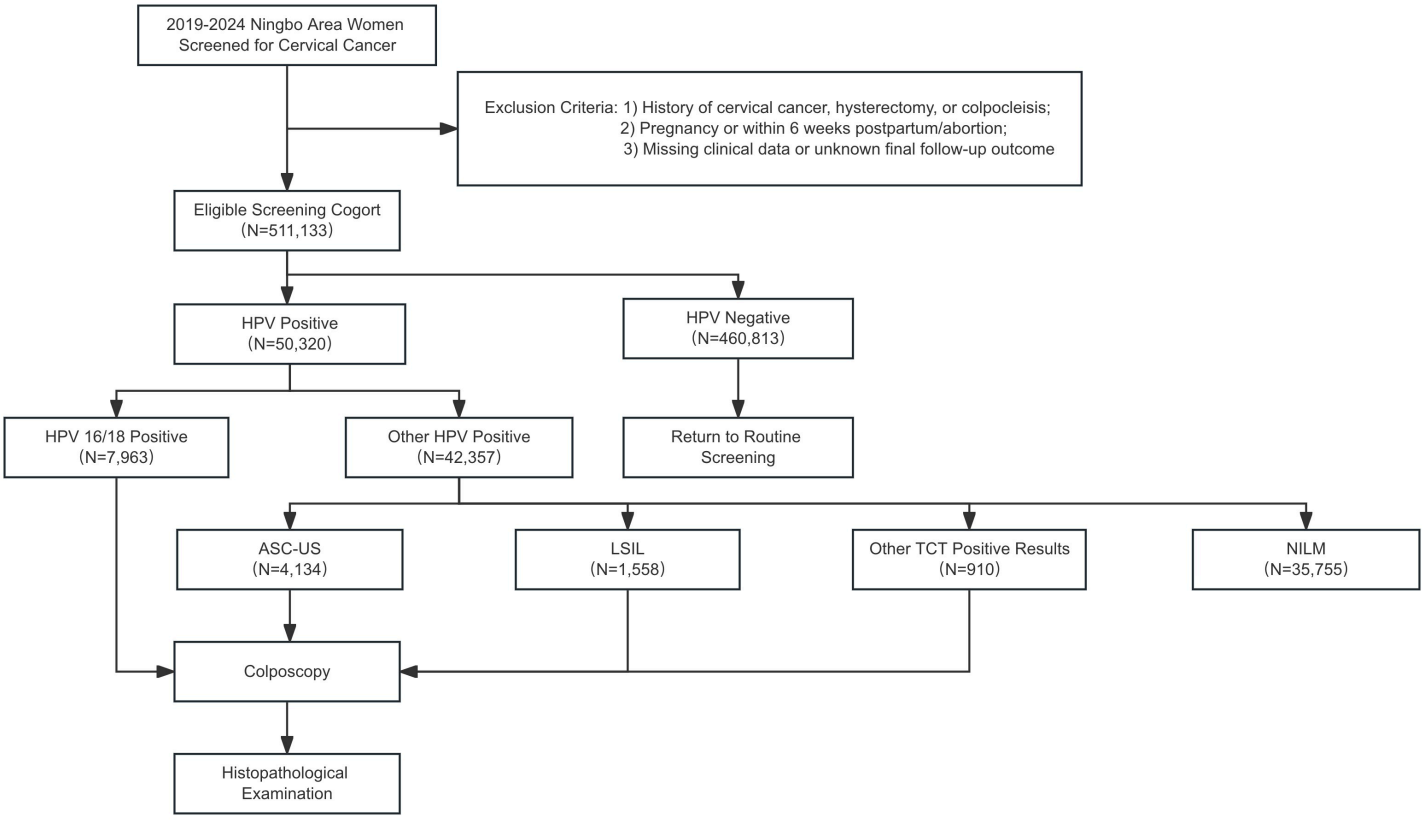

Figure 1

Clinical triage pathway and study flowchart of the cervical cancer screening population in Ningbo, China.

HPV: human papillomavirus; TCT: liquid-based thin-layer cytology; ASC-US: atypical squamous cells of undetermined significance; LSIL: low-grade squamous intraepithelial lesion; NILM: negative for intraepithelial lesion or malignancy.

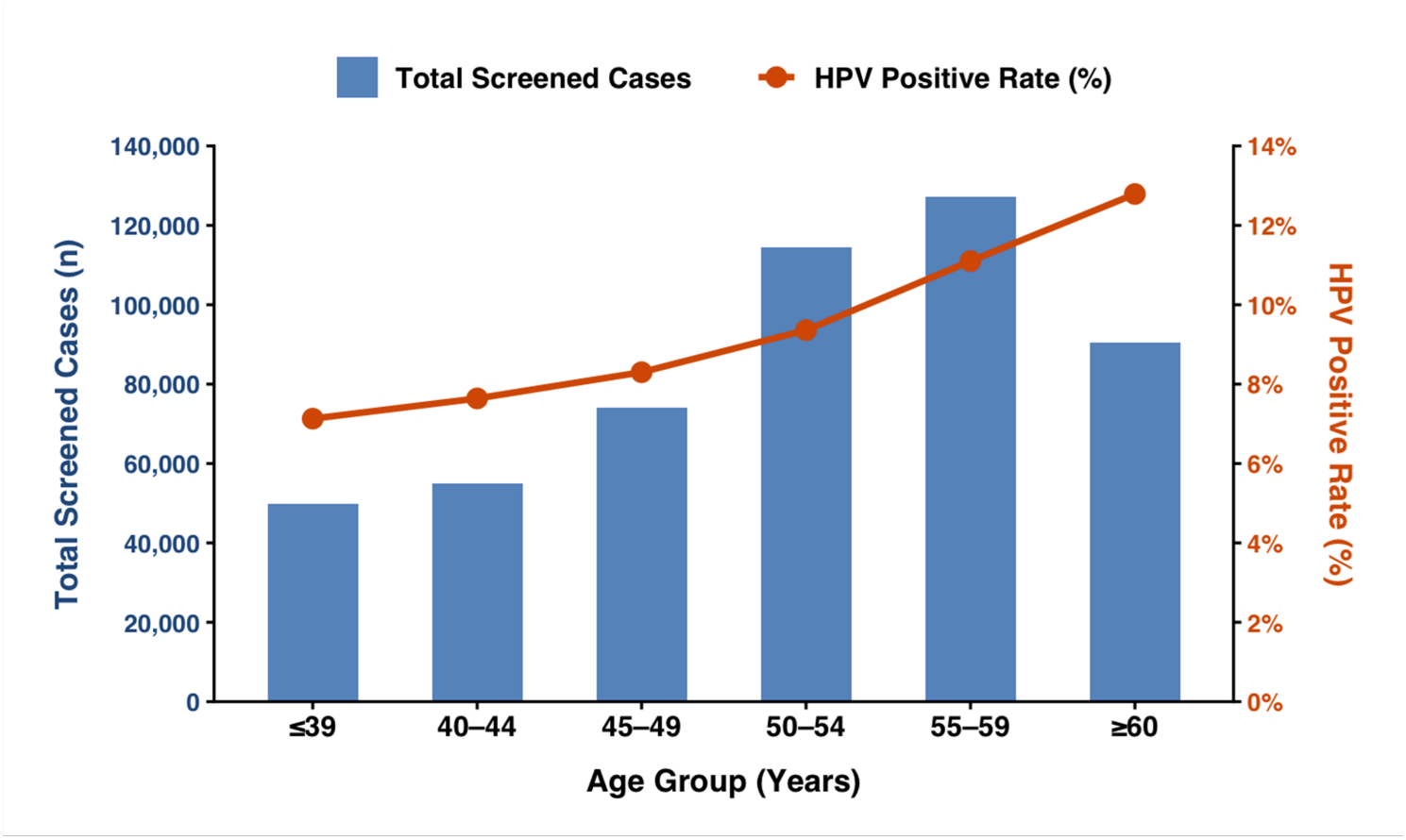

Figure 2

Age-specific distribution of the screened population and HPV positivity among 511,133 women undergoing cervical cancer screening in Ningbo, China.

The bars (left y-axis) represent the total number of women screened in each age group, and the line graph (right y-axis) depicts the corresponding HPV positivity rate (%). The data demonstrate a progressive upward trend in HPV positivity with advancing age, peaking in the ≥60-year-old group. HPV: human papillomavirus.

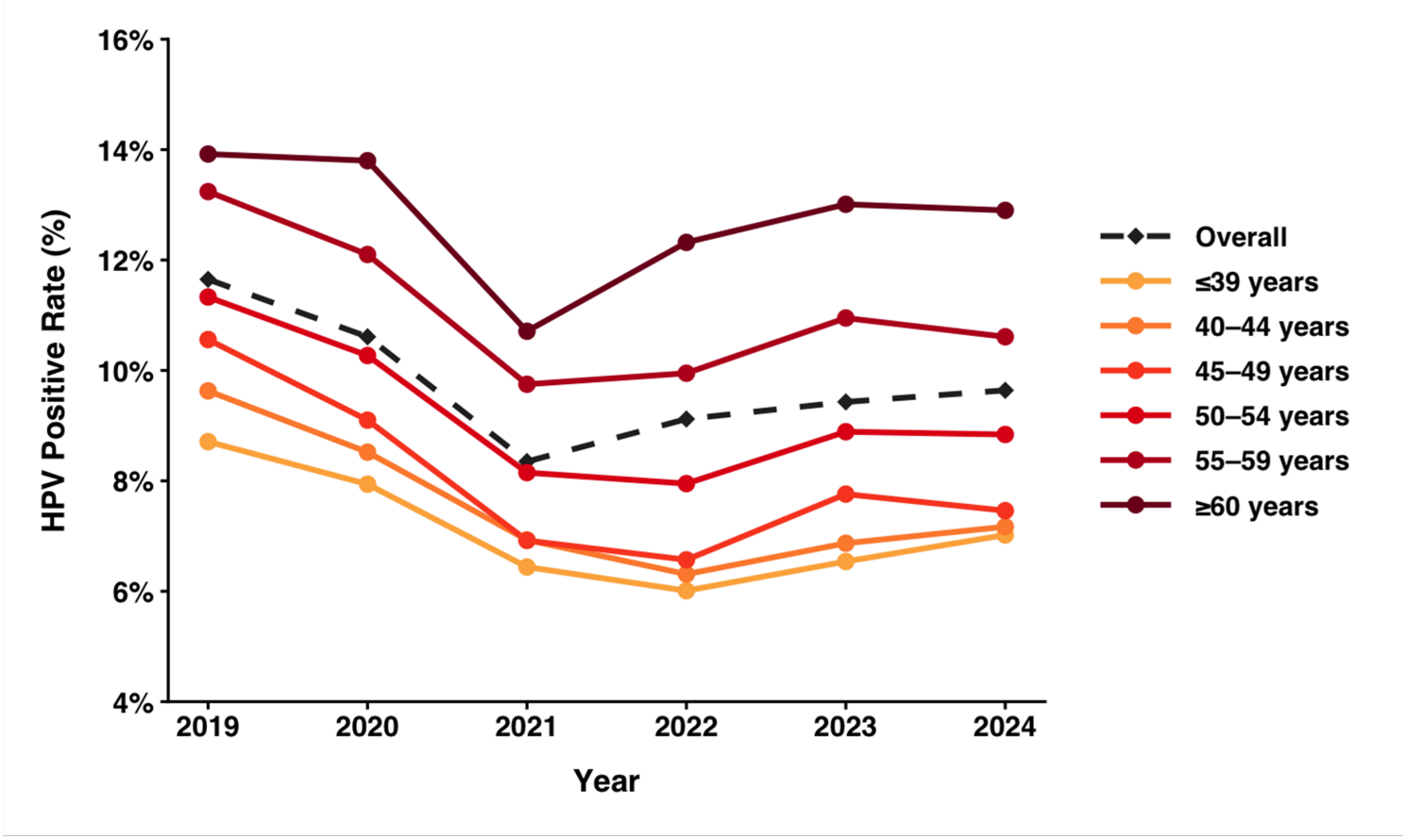

Figure 3

Dynamic temporal trajectories of age-specific HPV prevalence among women in Ningbo, China, from 2019 to 2024.

The line graph illustrates the six-year longitudinal changes in HPV infection rates (%) across six age brackets. Differently colored lines represent distinct age groups, with the color gradient (from light yellow to dark red) corresponding to advancing age (from the ≤39-year-old group to the ≥60-year-old group). Across all age strata, HPV prevalence reaches a trough during 2021-2022 before rebounding steadily. Throughout the six-year study period, HPV prevalence consistently exhibits an age-dependent upward pattern, with the ≥60-year-old group (dark red line) persistently bearing the highest infection burden. HPV: human papillomavirus.

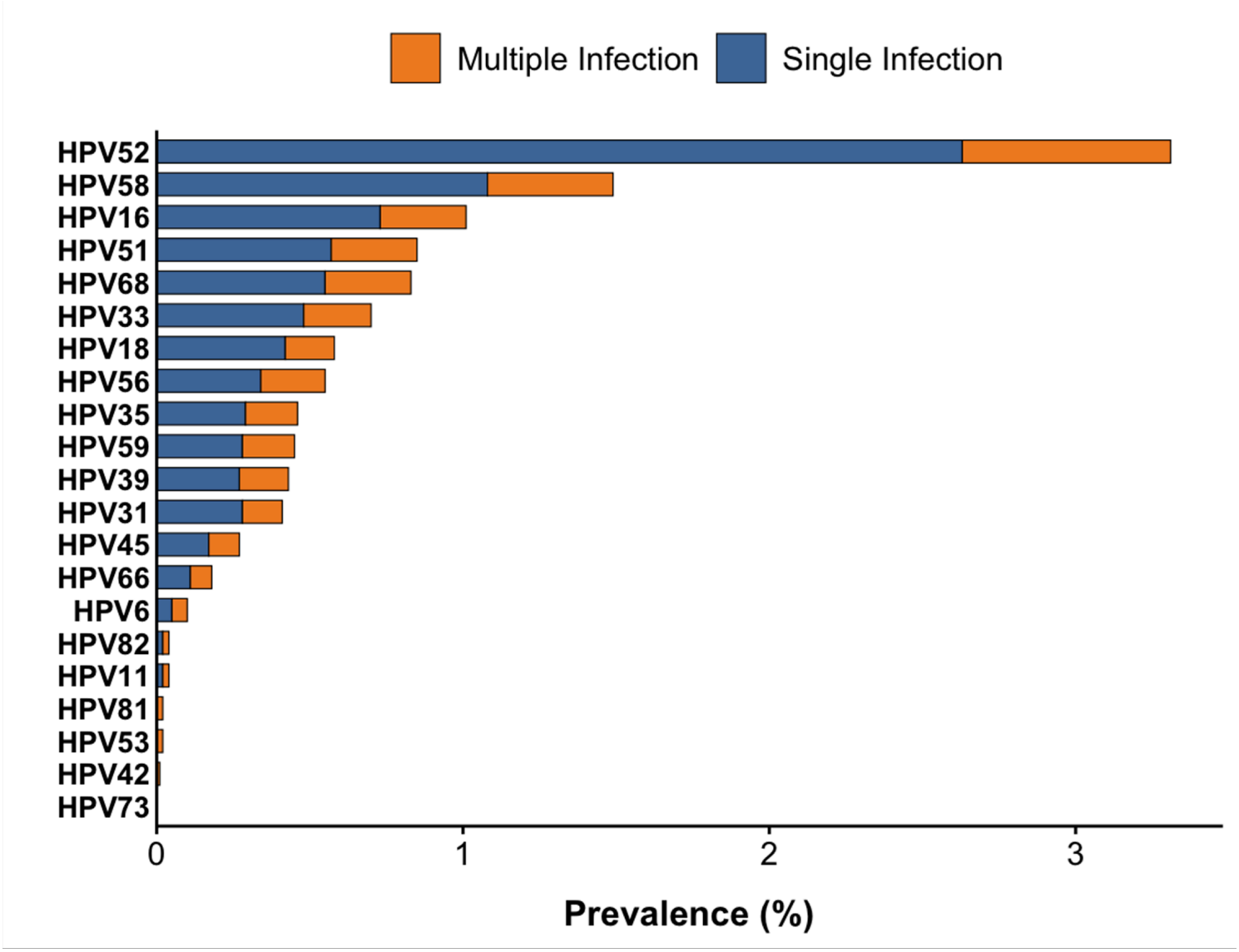

Figure 4

Prevalence and infection patterns of individual HPV genotypes among 511,133 women undergoing cervical cancer screening in Ningbo, China.

The horizontal stacked bar chart illustrates the overall prevalence (%) across 21 HPV genotypes, ranked in descending order by any-type estimate (Any.). The total length of each horizontal bar represents the overall prevalence burden of a specific genotype in the screened population, further subdivided into single-genotype infections (blue segments) and multiple infections (orange segments). The data indicate that HPV52, HPV58, and HPV16 are the top three predominant genotypes in Ningbo. Moreover, across the vast majority of genotypes—including the leading ones-single-genotype infections account for the predominant proportion of the overall infection burden. HPV: human papillomavirus; Any.: any-type estimate.

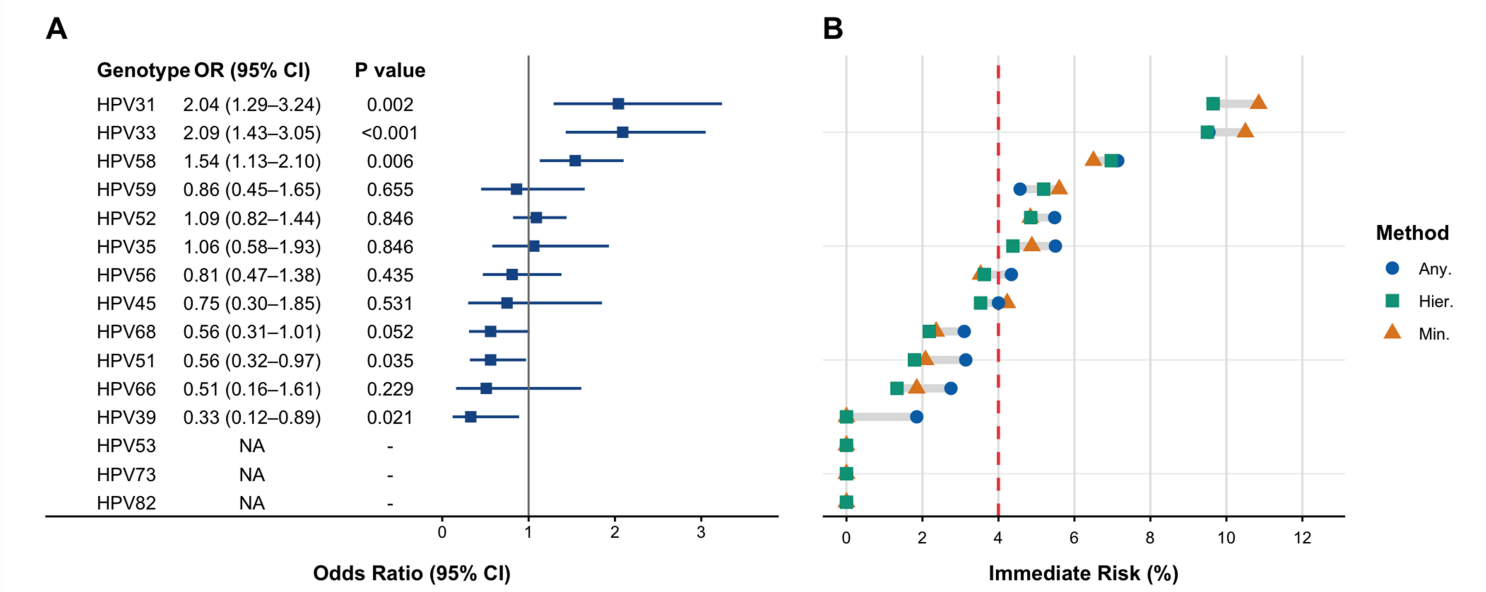

Figure 5

Immediate and relative risks of CIN2+ for non-16/18 hrHPV genotypes in an ASC-US background.

The figure illustrates the confounding effect of multiple infections on the risk assessment of individual genotypes. (A) Forest plot displaying the relative risks (ORs and 95% CIs) of CIN2+ driven by individual genotypes, using the average risk of the overall non-16/18-positive cohort as the baseline reference (OR = 1.0). The vertical solid line represents the null line (OR = 1.0). (B) Dumbbell plot showing the immediate CIN2+ risks (%) for each genotype in the ASC-US cohort, calculated via three statistical approaches. Blue circles, green squares, and orange triangles denote the immediate CIN2+ risks (%) under the Any., Hier., and Min. algorithms, respectively. The horizontal connecting line for each genotype represents the risk assessment span across different methods, visually reflecting the degree to which multiple infections interfere with the evaluation of a genotype's true independent risk. A wider span indicates a more pronounced confounding effect from multiple infections. Abbreviations: HPV, human papillomavirus; OR, odds ratio; CI, confidence interval.

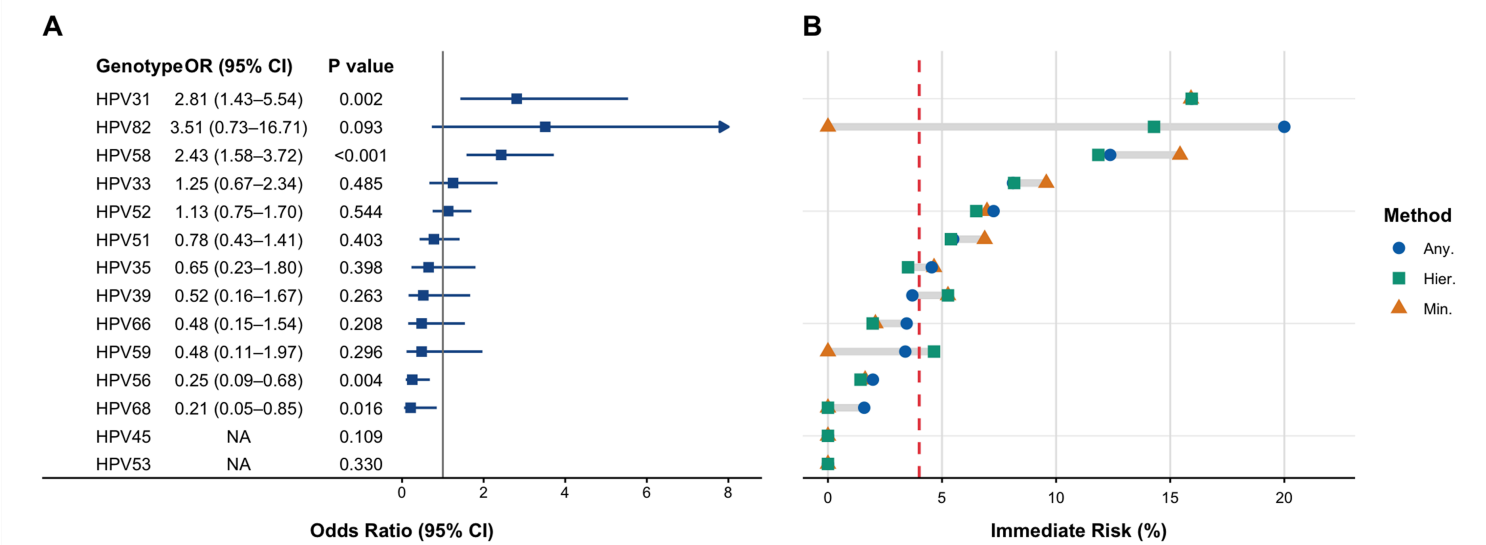

Figure 6

Immediate and relative risks of CIN2+ for non-16/18 hrHPV genotypes in an LSIL background.

The figure illustrates the confounding effect of multiple infections on the risk assessment of individual genotypes. (A) Forest plot displaying the relative risks (ORs and 95% CIs) of CIN2+ driven by individual genotypes, using the average risk of the overall non-16/18-positive cohort as the baseline reference (OR = 1.0). The vertical solid line represents the null line (OR = 1.0). (B) Dumbbell plot showing the immediate CIN2+ risks (%) for each genotype in the LSIL cohort, calculated via three statistical approaches. Blue circles, green squares, and orange triangles denote the immediate CIN2+ risks (%) under the Any., Hier., and Min. algorithms, respectively. The horizontal connecting line for each genotype represents the risk assessment span across different methods, visually reflecting the degree to which multiple infections interfere with the evaluation of a genotype's true independent risk. A wider span indicates a more pronounced confounding effect from multiple infections. Abbreviations: HPV, human papillomavirus; OR, odds ratio; CI, confidence interval.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.docx