Prevalence and type distribution of human papillomavirus in a Chinese urban population from 2014 to 2018: a retrospective study (#40939)

Second revision

Guidance from your Editor

Please submit by **23 Dec 2019** for the benefit of the authors .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Raw data check
Review the raw data. Download from the materials page.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
8 Figure file(s)
6 Table file(s)

Custom checks

Human participant/human tissue checks

- Have you checked the authors ethical approval statement?
- Does the study meet our article requirements?
- Has identifiable info been removed from all files?
- Were the experiments necessary and ethical?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Prevalence and type distribution of human papillomavirus in a Chinese urban population from 2014 to 2018: a retrospective study

Mei-Yan Xu [1], **Bing Cao** [2], **Yan Chen** [3], **Juan Du** [4], **Jian Yin** [1], **Lan Liu** [5], **Qing-Bin Lu** [Corresp. 4, 6]

[1] Department of Nutrition, Aerospace Center Hospital, Beijing, Beijing, China
[2] Key Laboratory of Cognition and Personality (SWU), Faculty of Psychology, Ministry of Education, Southwest University, Chongqing, Chongqing, China
[3] Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[4] Department of Laboratorial Science and Technology, School of Public Health, Peking University, Beijing, Beijing, China
[5] Department of Health Management, Aerospace Center Hospital, Beijing, Beijing, China
[6] Beijing Key Laboratory of Toxicological Research and Risk Assessment for Food Safety, Beijing, P.R. China, Beijing, Beijing, China

Corresponding Author: Qing-Bin Lu
Email address: qingbinlu@bjmu.edu.cn

Background Human papilloma virus (HPV) infection is the most common sexually transmitted infections among women worldwide. The current study's main objective is to report the prevalence and distribution of HPV types in an urban population in Beijing, China. **Methods** All the eligible female participants aged≥18 years were recruited from the Aerospace Center Hospital in Beijing, China between 2014 and 2018. A total of 21 HPV types were detected. **Results** In total, 12 high risk HPV types and nine low risk HPV types were detected. The HPV-positive rates were 8.85% in 2014, 7.16% in 2015, 7.60% in 2016, 8.31% in 2017, and 7.72% in 2018, respectively, in an urban population in Beijing, China. Overall, no significant differences in the HPV-positive rates were found over the five years. The peak prevalence of HPV infection in all types was observed in age group of 20-24 in all types. HPV52 was the dominant HPV type across the five years . Among all 21 HPV types, HPV66, HPV26, and HPV59 were ranked the top three in coinfection occurrence. **Conclusions** Our findings are great helpful for HPV screening and vaccination., and the associations between gynaecological diseases and the HPV types with high prevalence, particularly HPV52, warrant further investigation.

Prevalence and type distribution of human papillomavirus in a Chinese urban population between 2014 and 2018: a retrospective study

Mei-Yan Xu[1,#], Bing Cao[2,#], Yan Chen[3], Juan Du[4], Jian Yin[1], Lan Liu[5,*], Qing-Bin Lu[4,6*]

[1] Department of Nutrition, Aerospace Center Hospital, Beijing, P. R. China

[2] Key Laboratory of Cognition and Personality (SWU), Faculty of Psychology, Ministry of Education, Southwest University, Chongqing, P. R. China

[3] Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[4] Department of Laboratorial Science and Technology, School of Public Health, Peking University, Beijing, P. R. China

[5] Department of Health Management, Aerospace Center Hospital, Beijing, P. R. China

[6] Beijing Key Laboratory of Toxicological Research and Risk Assessment for Food Safety, Beijing, P.R. China

[#]Mei-Yan Xu and Bing Cao contributed equally.

Correspondence address

Qing-Bin Lu, M.D.,

Department of Laboratorial Science and Technology, School of Public Health, Peking University, Beijing, P.R. China.

Tel (fax): +86-10-82805327. Email: qingbinlu@bjmu.edu.cn

Lan Liu, Professor,

Department of Health Management, Aerospace Center Hospital, Beijing, P.R. China. Tel (fax): +86-10-59971711. Email: 721Lancy@sina.com

Running title

HPV prevalence

Abstract

Background

Human papilloma virus (HPV) infection is the most common sexually transmitted infections among women worldwide. The current study's main objective is to report the prevalence and distribution of HPV types in an urban population in Beijing, China.

Methods

All the eligible female participants aged≥18 years were recruited from the Aerospace Center Hospital in Beijing, China between 2014 and 2018. A total of 21 HPV types were detected.

Results

In total, 12 high risk HPV types and nine low risk HPV types were detected. The HPV-positive rates were 8.85% in 2014, 7.16% in 2015, 7.60% in 2016, 8.31% in 2017, and 7.72% in 2018, respectively, in an urban population in Beijing, China. Overall, no significant differences in the HPV-positive rates were found over the five years. The peak prevalence of HPV infection in all types was observed in age group of 20-24 in all types. HPV52 was the dominant HPV type across the five years. Among all 21 HPV types, HPV66, HPV26, and HPV59 were ranked the top three in coinfection occurrence.

Conclusions

Our findings are great helpful for HPV screening and vaccination., and the associations between gynaecological diseases and the HPV types with high prevalence, particularly HPV52, warrant further investigation.

Introduction

Human papilloma viruses (HPV) are small groups of double-stranded DNA viruses without enveloped icosahedral capsids (Boda et al. 2018; zur Hausen 2002) that cause the most common sexually transmitted infections among women worldwide (Handler et al. 2015; Satterwhite et al. 2013). More than 200 HPV types have been identified (Ghedira et al. 2016) and some had been confirmed to be associated with certain diseases, mainly cancers. Replicated evidence support that HPV infection causes of the development of most cervical, penile, vaginal, anal, vulvar, and oropharyngeal cancers and pre-cancers (Bansal et al. 2016). Epidemiological research has reported that 5.2% of all cancers worldwide can be attributed to HPV infection (Steben & Duarte-Franco 2007). Moreover, the overall annual direct medical cost burden of preventing and treating HPV-associated diseases was estimated as an annual $8.0 billion in the United States in 2010 (Chesson et al. 2012).

Each of the HPV type is identified with a number corresponding to the sequence of discovery. The International Agency for Research on Cancer (IARC) working group classified the following12 HPV types as carcinogenic to humans (Group 1) and high risk HPV (HR-HPV): including HPV16, HPV18, HPV31, HPV33, HPV35, HPV39, HPV45, HPV51, HPV52, HPV56, HPV58, and HPV59 (Clifford et al. 2011).

HPV type prevalence varies between different countries, but also within different regions of the same country (Wang et al. 2015). A meta-analysis reported that the estimated global HPV prevalence among women of all ages is 11.7%. The areas with the highest presences are Sub-Saharan Africa (24.0%), Eastern Europe (21.4%), and Latin America (16.1%), respectively (Bruni et al. 2010). The prevalence of HPV infections among mainland Chinese women is reported as 11.0% (Zhou et al. 2018), which is similar to the average level worldwide. Additional factors such as age, marital status, and underlying coinfection with other diseases or viruses such as human immunodeficiency virus (HIV), were reported to be associated with HPV infections (Shi et al. 2016).

HPV prevalence is still a concerning since there is no HPV vaccine that covers all HPV types associated with cervical cancer. Previous epidemiological studies have mainly focused on the prevalence of HR HPV types with specific cancers, and updated information on type-specific HPV prevalence and distribution in a general population is warranted. Therefore, we conducted a retrospective, hospital-based study to estimate the overall prevalence and distribution of HPV types in Beijing, China. The persistence rates of HPV infection after a 3-year follow up were also determined. The current findings will be valuable in the development and application of HPV vaccines in China.

Methods

Study design and participants

For this hospital-based retrospective epidemiological study, participants were recruited from the Aerospace Center Hospital located in Beijing, China between 2014 and 2018. The inclusion criteria of our subjects were as follows: 1) female, age 18 years old or older; 2) participated in annual health examinations and had HPV detected result at least once at the current hospital during our research period. We excluded 1) pregnant and postpartum women; and 2) the people who have been diagnosed with any kind of cancer. For our research samples, some had participated in annual health examinations every year, and some only participated once during our study period. The subject had a healthy examination once a year and had one test for HPV infection. We included all available data in the statistical analysis. We conducted a chart review to collect the recruited subjects' demographic information and physiological status details, including age and underlying diseases (such as diabetes or hypertension). Subjects were classified as having diabetes if they met the criteria for diabetes mellitus dictated in 2010 by the Japan Diabetes Society (Committee of the Japan Diabetes Society on the Diagnostic Criteria of Diabetes et al. 2010). Hypertension, by definition, is a repeatedly elevated blood pressure exceeding 140 over 90 mmHg (systolic pressure above 140 or a diastolic pressure above 90).

The research protocol was approved by the human ethics committee of the Aerospace Center Hospital and all methods were carried out in accordance with the approved guidelines (No. 2014YN-01). All participants provided written or verbal informed consent to have their samples and information collected and used for this study.

Outcome measurement

Cervicovaginal cellular swabs were collected by a doctor during gynecological healthy examination. Most samples were tested on the same day after collection; otherwise, they were stored at -70°C until tested.

Viral DNA was isolated from cervicovaginal cellular swabs using a QIAamp DNA Mini Kit (QIAGEN). In the earlier years, the 15HPV types were detected using the polymerase chain reaction (PCR) reverse dot blot method with the HPV Genotyping Kit [Yaneng Biotechnology (Shenzhen) Co., Ltd, Shenzhen, China], which enabled the detection of 12 HR-HPV types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59) and three low risk HPV (LR-HPV) types (HPV6, 11 and 68). An aliquot of a 5 µL DNA sample was briefly used, and the PCR reaction was amplified in a thermal cycler under the following conditions: an initial 15 min at 50°C, 10 min at 95°C; 40 cycles of 30 s at 94°C, 90 s at 42°C, and 30 s at 72°C; and a final extension 5 min at 5 min. The PCR products were immobilized onto a nitrocellulose membrane and hybridized with typing probes. Final results were judged by the direct visualization of the location of blue spots located on the membrane. In recent years, 21 HPV types were detected using the fluorescence quantitative PCR method with a nucleic acid detection kit (Jiangsu Shuoshi Biotechnology Co., Ltd, Taizhou, China) that enables the detection of 12 HR-HPV types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59) and nine LR-HPV types (HPV6, 11, 26, 53, 58, 66, 73, 81 and 82; HPV26, 53, 66, 73 and 82 were considered as LR-HPV types in this study). Each PCR reaction was performed briefly in a 40 μL mixture containing 4 μL of extracted DNA and 36 μL of PCR master mix, at 94°C for 2 min, 40 cycles of 10 s at 93°C, and 30 s at 62°C. Specific primers and correspondent fluorescent probes were designed to detect 21 HPV types in the assay. A single copy gene was amplified to serve as an internal quality control for DNA preparation. A sample was determined as HPV positive when the cycle threshold (t value) was less than or equal to 38.0, and the amplification curve was a typical S-type. According to the preliminary experiments in the manuals, the sensitivity and specificity in both the two kits were more than 98% (http://www.yanengbio.com/HPV/ and http://www.s-sbio.com/products_detail/productId=512.html). Detailed information on the two detection methods is listed in the Supplemental Table 1.

Statistical analysis

Annual HPV prevalence rates and the distribution of HPV types in each year were calculated using positive rates or proportions and 95% confidence intervals (CIs), and they were stratified by age at diagnosis. We explored the differences of HPV detection across various demographics (i.e., age) and clinical variables (i.e., diabetes, hypertension) using either an independent t-test or χ^2 test, respectively. A positive HPV sample was defined as having an infection of HPV types equal or more than one. A single infection was defined as being infected by one type of HPV infection. The coinfection proportion was defined as the number of subjects infected with more than one HPV type divided by the total number of subjects infected with at least one HPV type. The persistence rate of HPV infection was calculated as (①+ and ②+)/(①+ and ②) for one year, (①+ and ②+ and ③+)/(①+ and ② and ③) for two years and (①+ and ②+ and ③+ and ④+)/(①+ and ② and ③ and ④) for three years, where ①, ②, ③, and ④ mean the detection at the first, second, third and fourth year, respectively. A two-sided p value of less than 0.05 was considered significant. All statistical analyses were done using Stata 14.0 (Stata Corp LP, College Station, TX, USA).

Results

Overall prevalence and basic characteristic distribution of HPV infection

Between January 1, 2014 and August 31, 2018, 5880 women involved in this study had at least one HPV detection (Supplemental Fig. 1). The annual age distributions of the participants every year are shown in Supplemental Fig. 2. The age distribution was similar across every year and the majority of participants were between 30-45 years old. There were 837 (14.2%) subjects with diabetes and 1491 (25.4%) subjects with hypertension among all 5880 female participants. As shown in Fig.1, the HPV positive rates were 8.85% in 2014, 7.16% in 2015, 7.60% in 2016, 8.31% in 2017, and 7.72% in 2018, respectively. No significant differences in HPV positive rates were found over the five years (p = 0.357).

The basic participant characteristics between HPV (+) and HPV (-) cases are ~~were~~ presented in Table 1. No significant differences were found in the distribution of diabetes between HPV (+) and HPV (-) across all of the five years. In terms of age distribution, the participants in the HPV (+) group were generally comparable with those in the HPV (-) group. Lower hypertension rates were found in the HPV (+) group in 2015 and in 2017, with $p = 0.021$ and $p = 0.009$, respectively.

HPV type distribution

Twelve HR-HPV and nine LR-HPV types were detected, and a significant difference was found in the LR-HPV positive rate between 2014 and 2018 ($p < 0.001$). In contrast, the HR-HPV rate decreased over time (Fig. 2A). The HPV positive rates varied across different age groups. Here we identified the age distribution of HPV infection. As shown in Fig. 2B, most HPV infections occurred in age groups of 20-24 years old, 65-69 years old, and 50-54 years old. Young women aged 20-24 years had the highest LR-HPV and HR-HPV infection rates (Fig. 2C and 2D), and the detection rates of HPV06 and HPV11 also peaked in the 20-24 years old age group (Fig. 2E).

The positive rate distributions of all 21 detected HPV types were different in each year (Fig. 3 and Supplemental Fig. 3). The most dominant HPV type was HPV52, which occupied the top position all five years. HPV58, HPV16, and HPV53 were also ranked high but their order slightly changed each year. An interesting phenomenon in 2014 was that the positive rate of some HPV subtypes (especially LR-HPV) was very low or even zero, although their infection rates increased every year.

Coinfection and persistence of HPV types infection

Multiple HPV infections were very common, and the ranking of HPV type by coinfection status is shown in Fig. 4. About 75% of the HPV types had a coinfection proportion ≥ 50%. Among all 21 HPV types, HPV66, HPV26, and HPV59 were ranked the highest in coinfections. HPV52 had the most positive rate but was ranked the lowest. The distribution of double HPV coinfections is listed in Supplemental Table 2.

Table 2 shows the HPV positive rates in the follow-up periods after the first HPV positive test. Significant differences in HPV positive rates were found in both the persistent infection group and the non-persistent infection group during the 3-years follow-up ($p < 0.001$). Persistent infections of different HPV types in four different years were detected in 16 cases. As shown in Supplemental Fig. 4, HPV52, HPV58, HPV35, and HPV68 were more likely to cause persistent infections.

Discussion

The main objective of this retrospective study is to report the prevalence and distribution of HPV types based on the annual health examinations and HPV detection results of women aged ≥ 18 years between 2014 and 2018 at the Aerospace Center Hospital in Beijing, China. The second objective of the study is to contribute more information and contribute to protect women from further HPV infections and related cancers in China. Most previous studies mainly focused on HR-HPV types (Guardado-Estrada et al. 2014; Nielsen et al. 2009). To the best of our knowledge, this is the first large-scaled study to provide the prevalence, distribution, and persistence of both HR-HPV and LR-HPV types in an urban population in China.

The prevalence rate of HPV infection revealed no changes over the years, which means that HPV infection exists steadily and persistently, sustainably damaging women's health. The prevalence rate of HPV infection can be controlled and reduced through the use of HPV vaccines, making the promotion and acceptance of HPV vaccines crucial.

The prevalence rate of HR HPV types shows a slightly decreasing trend overall in all ages over the five years between 2014 and 2018. The prevalence rate of LR HPV, however, shows an overall increasing trend. Therefore, the variety of HPV infections were increased, although the prevalence rate of HPV infection was reduced. In regard to the distribution of HPV infections across different age groups, the peak prevalence of all HR-HPV and LR-HPV types of HPV infections was observed in the 20-24 years old age group. The prevalence of LR-HPV types in women aged 20-24 years is close to 6%, which is significantly high compared to less than 2% observed in other age groups. Women aged 65-69 years also showed significantly high prevalence in all HR types. However, these results may be biased by the different detection methods used in this study.

Because of the large population base of China, there is imbalanced development and inconsistent allocations of healthcare resources and public health awareness (Qiao 2018). A previous study showed that women aged 40-49 years were the peak age group for HPV infection prevalence, accounting for 17.9 % of cases in in Ningbo, south-eastern China (Baloch et al. 2017). According to a study performed in Yunnan, south-western China, women with the highest HPV prevalence were ≤ 29 years old in the urban regions and ≥ 50 years in rural regions. These studies did not achieve uniform results on age distribution (Baloch et al. 2017). In addition to age, location (i.e. rural/ urban areas) and healthcare resources, other risk factors may contribute to the prevalence and distribution of HPV, such as marital status, sexual activity, genetic variants, and coinfections with other diseases (i.e. HIV) (Zhao et al. 2018), and education, public health awareness, alcohol use, and tobacco use may be potential confounding variables (Baloch et al. 2017).

Over the five-year period, HPV52 was the most dominant type among patients with positive detections, supporting the findings of several other studies (Baloch et al. 2017; Hong et al. 2015). It is followed by HPV58 and HPV16 (Fig.2) as the most common types in female residents in Shanghai (Xu et al. 2018). In addition to single infections caused by HPV types, HPV52 also accounted for over 80% of coinfection cases, followed by HPV58 and HPV16. HPV26, HPV73 and HPV82 had the lowest single infection/coinfection proportions. The high prevalence of HPV52 in our study population and its association with other gynecological diseases warrant further investigation.

According to the previously reported data, over 70% of women have had an HPV infection at least once in their lifetime, with about 10% having a lifelong persistent infection (Liao et al. 2018). Persistent HPV infection can cause ~~the destructions of epithelial cells, the~~ abnormal proliferation of cells, and the accumulations of genetic damage leading to cancer of the cervix, vulva, vagina, anus, penis, and/or oropharynx (Crosbie et al. 2013; Hong et al. 2015). It is the second most common cause of death from cancer for women globally (Lowy & Schiller 2006), with higher death rates reported in lower income countries (Qiao 2018). According to our findings, persistent infection has a statistically significant effect on the outcome of ~~positive~~ HPV-positive detection over a three-year follow-up. High-risk HPV16 and HPV18 account for over 70% of cases of cervical cancer, which is the fourth most common cancer among women worldwide (Qiao 2018), while LR-HPV6 and LR-HPV11 account for about 90% of external genital warts cases (Lowy & Schiller 2006).

Prevention methods such as screening, early-detection and treatment, and vaccination are used to fight HPV viruses(Qiao 2018). Three types of vaccinations are available in mainland China: the bivalent vaccine which targets HPV16 and HPV18; the tetravalent vaccine which targets HPV6, HPV11, HPV16 and HPV18; and the nine-valent (9v) vaccine targets five additional HR-HPV types (HPV31, HPV33, HPV45, HPV52 and HPV58) (Crosbie et al. 2013; Lowy & Schiller 2006). The 9v vaccine (Gardasil 9) was approved by the Food and Drug Administration of China and was recently released into the market on April. 28, 2018. A set of three vaccines costs about 5,800 CNY, which is about twice the price in North America countries (i.e. America, Canada). Due to limited quantities imported from the United States to a high number of clinics and hospitals, vaccines are scare in China. Women 25 years and older may be able to afford the cost and long wait time, but younger women may not be able to afford the vaccines themselves if their healthcare programs do not cover the cost. Implementing HPV prevention initiatives may not meet the current need (Ogembo et al. 2015).

All data in the current study are based on highly reliable hospital records. The large sample size-based population means the results of this study are generalizable. Limited data are available on the prevalence and type distribution of HPV infections in Beijing, the capital of China making the present study valuable for the future clinical researches of HPV screening and anti-HPV vaccines. HPV distribution largely varies and is affected by many confounding variables. The results and interpretations presented in the current study should be considered in the context of the following limitations. First, all participants were recruited from the Aerospace Center Hospital in Beijing and most of them were from urban areas, which is not sufficiently representative of the whole population and may be selection biased. Second, since individuals in China can choose their type of physical examination from a variety of packages and can voluntarily choose or decline HPV type testing, a large number of people do not have the results from all five years. Additionally, we can only perform analysis of persistent infection in a small number of participants with follow-up test results. A detection bias for some HPV types exists in the study due to changing and unbalanced distributions of detection methods, especially for the LR HPV types, which may be related to changes in LR or HR HPV types.

Conclusion

The current retrospective study presents the prevalence and distribution of HPV types in women between 2014 and 2018 in an urban population of Beijing, China. HPV52 was the most dominant HPV type over all five years. The peak prevalence of all HPV infection types was observed in ages 20-24 years old in all types. The current findings can be significant for the future of HPV screening and vaccination. The associations between gynaecological diseases and HPV types with high prevalence, especially HPV52, warrant further investigation.

Declarations

Ethics approval and consent to participate

The research protocol was approved by the human ethics committee of the Aerospace Center Hospital and all methods were carried out in accordance with the approved guidelines (No. 2014YN-01). All participants provided written or verbal informed consent to have their samples and information collected and used for this study.

Competing interests

The authors declare that they have no competing interests.

Funding

The study was supported by China Mega-Project for Infectious Diseases Grant (2017ZX10103004), the Youth Innovation Funding (2014QN09) and National Natural Science Foundation of China (No. 81703274) and Peking University Medicine Seed Fund for Interdisciplinary Research (BMU2018MX009). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

Conceived and designed the experiments: QB Lu, MY Xu, L Liu and B Cao. Collect data and performed the experiments: MY Xu, J Yin and L Liu. Analyzed the data: QB Lu, MY Xu and B Cao. Contributed reagents/materials/analysis tools: QB Lu, J Du, S Wang, Y Chen, MY Xu and L Liu. Wrote the paper: QB Lu, MY Xu, B Cao, Y Chen and L Liu.

Acknowledgements

None.

References

Baloch Z, Yasmeen N, Li Y, Ma K, Wu X, Yang SH, and Xia X. 2017. Prevalence and risk factors for human papillomavirus infection among Chinese ethnic women in southern of Yunnan, China. *Brazlilian Journal of Infectious Diseases* 21:325-332. 10.1016/j.bjid.2017.01.009

Bansal A, Singh MP, and Rai B. 2016. Human papillomavirus-associated cancers: A growing global problem. *International Journal of Applied and Basic Medical Research* 6:84-89. 10.4103/2229-516X.179027

Boda D, Docea AO, Calina D, Ilie MA, Caruntu C, Zurac S, Neagu M, Constantin C, Branisteanu DE, Voiculescu V, Mamoulakis C, Tzanakakis G, Spandidos DA, Drakoulis N, and Tsatsakis AM. 2018. Human papilloma virus: Apprehending the link with carcinogenesis and unveiling new research avenues (Review). *International Journal of Oncology* 52:637-655. 10.3892/ijo.2018.4256

Bruni L, Diaz M, Castellsague X, Ferrer E, Bosch FX, and de Sanjose S. 2010. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *The Journal of Infectious Diseases* 202:1789-1799. 10.1086/657321

Chesson HW, Ekwueme DU, Saraiya M, Watson M, Lowy DR, and Markowitz LE. 2012. Estimates of the annual direct medical costs of the prevention and treatment of disease associated with human papillomavirus in the United States. *Vaccine* 30:6016-6019. 10.1016/j.vaccine.2012.07.056

Clifford GM, Howell-Jones R, and Franceschi S. 2011. Judging the carcinogenicity of human papillomavirus types by single/multiple infection ratio in cervical cancer. *International Journal of Cancer* 129:1792-1794. 10.1002/ijc.25833

Committee of the Japan Diabetes Society on the Diagnostic Criteria of Diabetes M, Seino Y, Nanjo K, Tajima N, Kadowaki T, Kashiwagi A, Araki E, Ito C, Inagaki N, Iwamoto Y, Kasuga M, Hanafusa T, Haneda M, and Ueki K. 2010. Report of the Committee on the Classification and Diagnostic Criteria of Diabetes Mellitus. *Journal of Diabetes Investigation* 1:218-222.

Crosbie EJ, Einstein MH, Franceschi S, and Kitchener HC. 2013. Human papillomavirus and cervical cancer. *The Lancet* 382:889-899. 10.1016/S0140-6736(13)60022-7

Ghedira R, Mahfoudh W, Hadhri S, Gabbouj S, Bouanene I, Khairi H, Chaieb A, Khelifa R, Bouaouina N, Remadi S, Elmi AA, Bansal D, Sultan AA, Faleh R, Zakhama A, Chouchane L, and Hassen E. 2016. Human papillomavirus genotypes and HPV-16 variants distribution among Tunisian women with normal cytology and squamous intraepithelial lesions. *Infectious Agents and Cancer* 11:61. 10.1186/s13027-016-0109-2

Guardado-Estrada M, Juarez-Torres E, Roman-Bassaure E, Medina-Martinez I, Alfaro A, Benuto RE, Dean M, Villegas-Sepulveda N, and Berumen J. 2014. The distribution of high-risk human papillomaviruses is different in young and old patients with cervical cancer. *PLoS One* 9:e109406. 10.1371/journal.pone.0109406

Handler MZ, Handler NS, Majewski S, and Schwartz RA. 2015. Human papillomavirus vaccine trials and tribulations: Clinical perspectives. *Journal of the American Academy of Dermatology* 73:743-756; quiz 757-748. 10.1016/j.jaad.2015.05.040

Hong H, He TF, Ni HX, Zhang S, and Xu GZ. 2015. Prevalence and genotype distribution of HPV infection among women in Ningbo, China. *International Journal of Gynecology & Obstetrics* 131:96-99. 10.1016/j.ijgo.2015.04.027

Liao L, Cheng H, Zeng F, Zhou W, and Ding Y. 2018. Prevalence and distribution of human papillomavirus genotypes among women with high-grade squamous intraepithelial lesion and invasive cervical cancer in Ganzhou, China. *Journal of Clinical Laboratory Analysis*:e22708. 10.1002/jcla.22708

Lowy DR, and Schiller JT. 2006. Prophylactic human papillomavirus vaccines. *Journal of Clinical Investigation* 116:1167-1173. 10.1172/JCI28607

Nielsen A, Iftner T, Munk C, and Kjaer SK. 2009. Acquisition of high-risk human papillomavirus infection in a population-based cohort of Danish women. *Sexually Transmitted Infections* 36:609-615. 10.1097/OLQ.0b013e3181a96d0e
Ogembo RK, Gona PN, Seymour AJ, Park HS, Bain PA, Maranda L, and Ogembo JG. 2015. Prevalence of human papillomavirus genotypes among African women with normal cervical cytology and neoplasia: a systematic review and meta-analysis. *PLoS One* 10:e0122488. 10.1371/journal.pone.0122488
Qiao YL. 2018. A long journey of cervical cancer prevention in China. *Chinese Journal of Oncology* 40:721-723. 10.3760/cma.j.issn.0253-3766.2018.10.001
Satterwhite CL, Torrone E, Meites E, Dunne EF, Mahajan R, Ocfemia MC, Su J, Xu F, and Weinstock H. 2013. Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2008. *Sexually Transmitted Infections* 40:187-193. 10.1097/OLQ.0b013e318286bb53
Shi Y, Peng SL, Yang LF, Chen X, Tao YG, and Cao Y. 2016. Co-infection of Epstein-Barr virus and human papillomavirus in human tumorigenesis. *Chinese Journal of Cancer* 35:16. 10.1186/s40880-016-0079-1
Steben M, and Duarte-Franco E. 2007. Human papillomavirus infection: epidemiology and pathophysiology. *Gynecologic Oncology* 107:S2-5. 10.1016/j.ygyno.2007.07.067
Wang R, Guo XL, Wisman GB, Schuuring E, Wang WF, Zeng ZY, Zhu H, and Wu SW. 2015. Nationwide prevalence of human papillomavirus infection and viral genotype distribution in 37 cities in China. *BMC Infectious Diseases* 15:257. 10.1186/s12879-015-0998-5
Xu J, Xia Z, Wang L, Yang B, Zhu Y, Zhu X, and Xu L. 2018. Human papillomaviruses 16 and 58 are distributed widely among women living in Shanghai, China, with high-grade, squamous intraepithelial lesions. *Epidemiology and Infection*:1-6. 10.1017/S0950268818003011
Zhao P, Liu S, Zhong Z, Hou J, Lin L, Weng R, Su L, Lei N, Hou T, and Yang H. 2018. Prevalence and genotype distribution of human papillomavirus infection among women in northeastern Guangdong Province of China. *BMC Infectious Diseases* 18:204. 10.1186/s12879-018-3105-x
Zhou HL, Zhang W, Zhang CJ, Wang SM, Duan YC, Wang JX, Yang H, and Wang XY. 2018. Prevalence and distribution of human papillomavirus genotypes in Chinese women between 1991 and 2016: A systematic review. *Journal of Infection* 76:522-528. 10.1016/j.jinf.2018.02.008
zur Hausen H. 2002. Papillomaviruses and cancer: from basic studies to clinical application. *Nature Reviews Cancer* 2:342-350. 10.1038/nrc798

Figure legends

Fig. 1 The prevalence rate of HPV from 2014 to 2018.

Fig. 2 The prevalence rate of low-risk and high-risk HPV types in different age groups. (A) Prevalence rate of low-risk and high-risk HPV types in each year; (B) prevalence of all HPV types of different age groups; (C) prevalence of low-risk HPV types of different age groups; (D) prevalence of high-risk HPV types of different age groups; (E) prevalence of HPV06 and HPV11.

Fig. 3 The annual distributions of every HPV types from 2014 to 2018.

Fig. 4 Multiple HPV infections and the ranking of coinfection status of each HPV type.

Supplemental Fig. 1 The diagram flowchart of the participants in this study.

Supplemental Fig. 2 The number of participants in each age groups during 2014-2018.

Supplemental Fig. 3 The annual distributions of every HPV types in each year from 2014 to 2018.

Supplemental Fig. 4 The persistent infections of different HPV types in four different years detected in 16 participants.

Figure 1

The prevalence rate of HPV from 2014 to 2018

The prevalence rate of HPV from 2014 to 2018

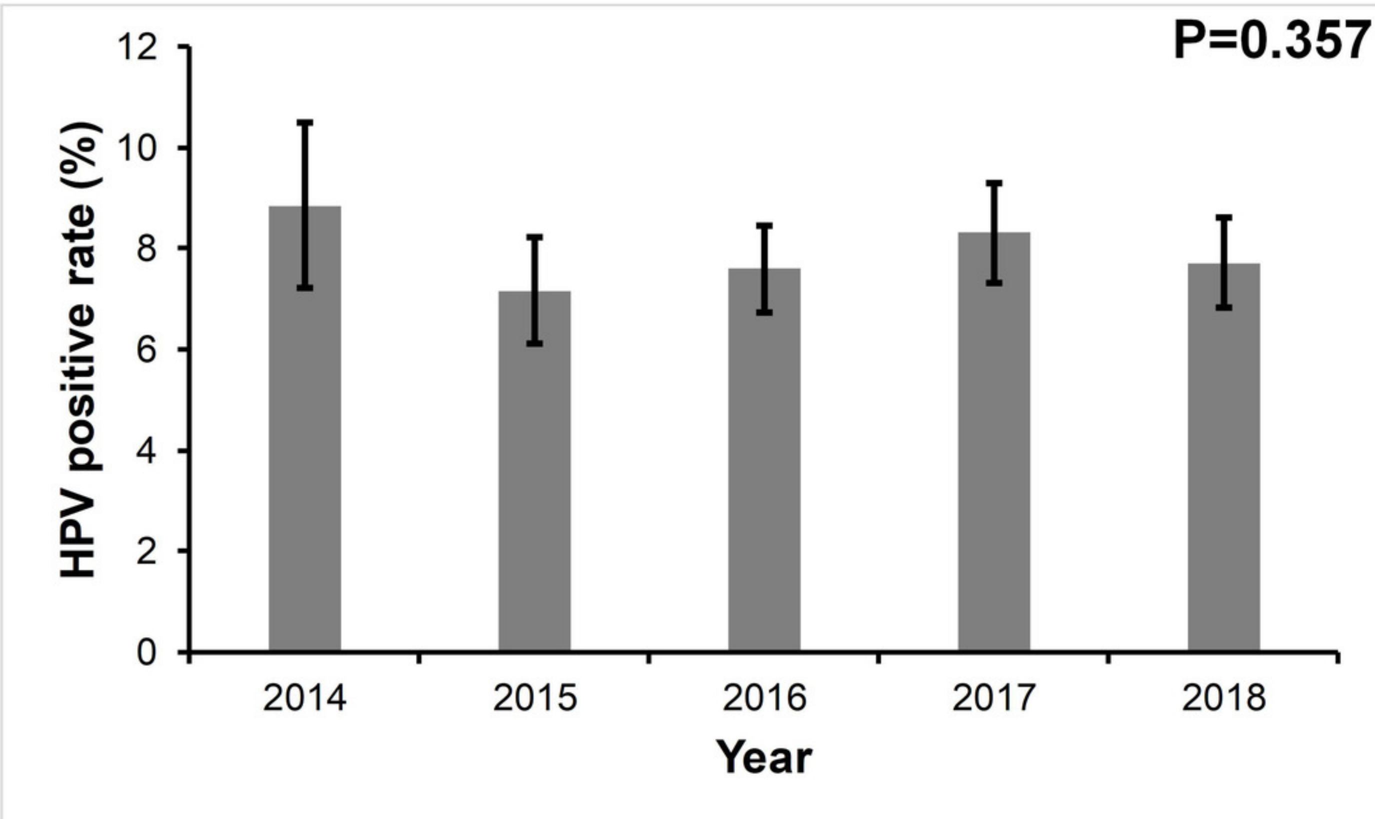

Figure 2

The prevalence rate of low risk and high risk HPV genotypes in different age groups.

(A) prevalence rate of low risk and high risk HPV genotypes in each year; (B) prevalence of all HPV genotypes of different age groups; (C) prevalence of low risk HPV genotypes of different age groups; (D) prevalence of high risk HPV genotypes of different age groups.

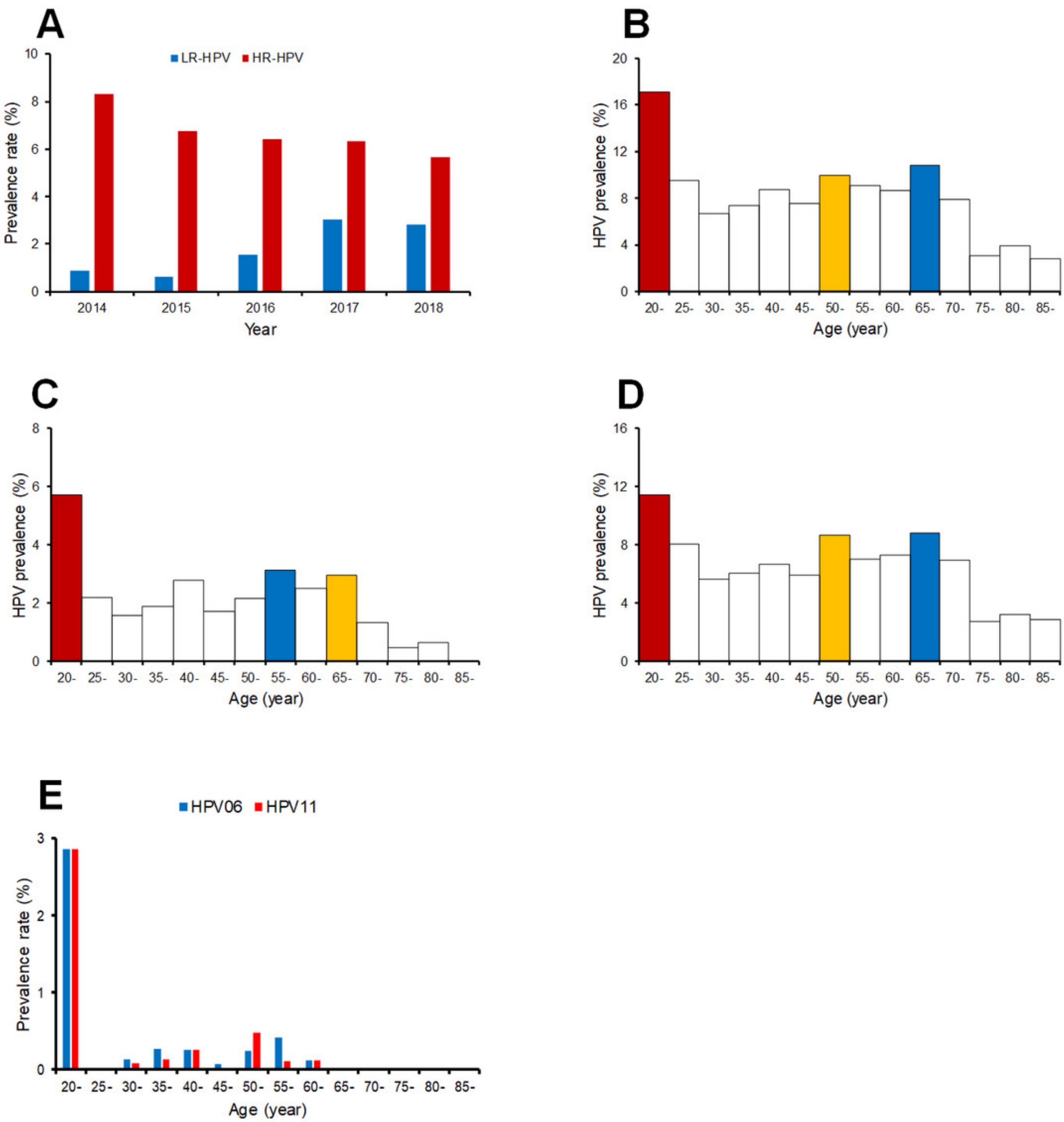
A
LR-HPV
HR-HPV
Prevalence rate (%)
Year
B
HPV prevalence (%)
Age (year)
C
HPV prevalence (%)
Age (year)
D
HPV prevalence (%)
Age (year)
E
HPV06
HPV11
Prevalence rate (%)
Age (year)

Figure 3

The distributions of every HPV genotypes in each year from 2014 to 2018.

The distributions of every HPV genotypes in each year from 2014 to 2018.

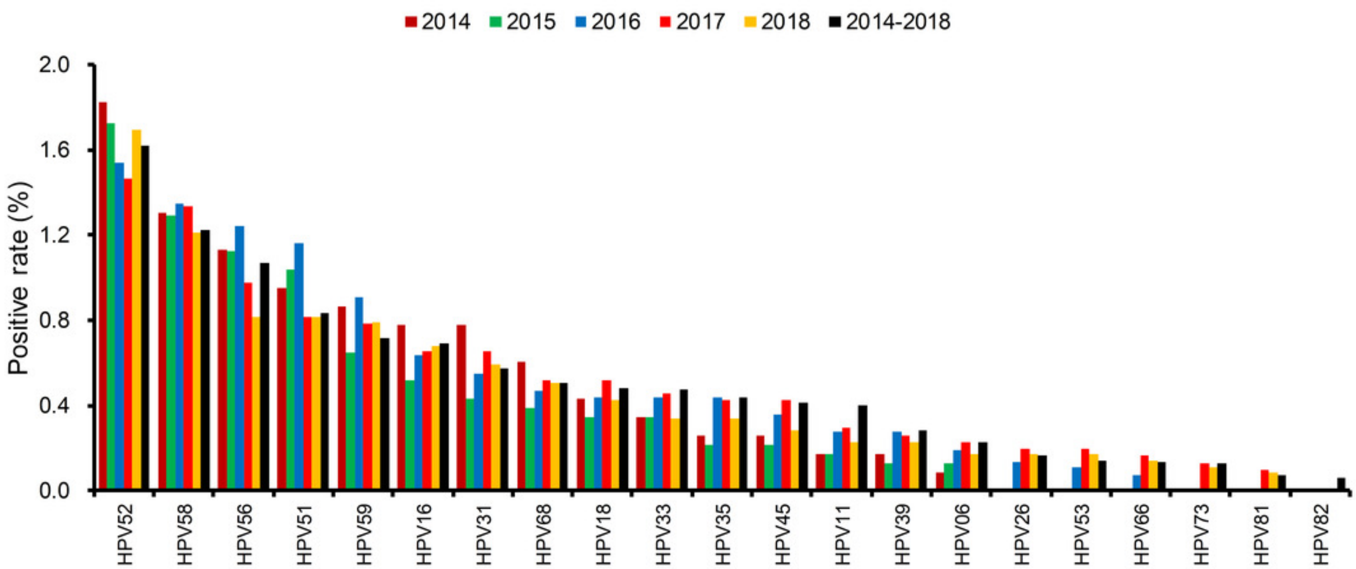

Figure 4

Multiple HPV infections and the ranking of coinfection status of each HPV genotype.

Multiple HPV infections and the ranking of coinfection status of each HPV genotype.

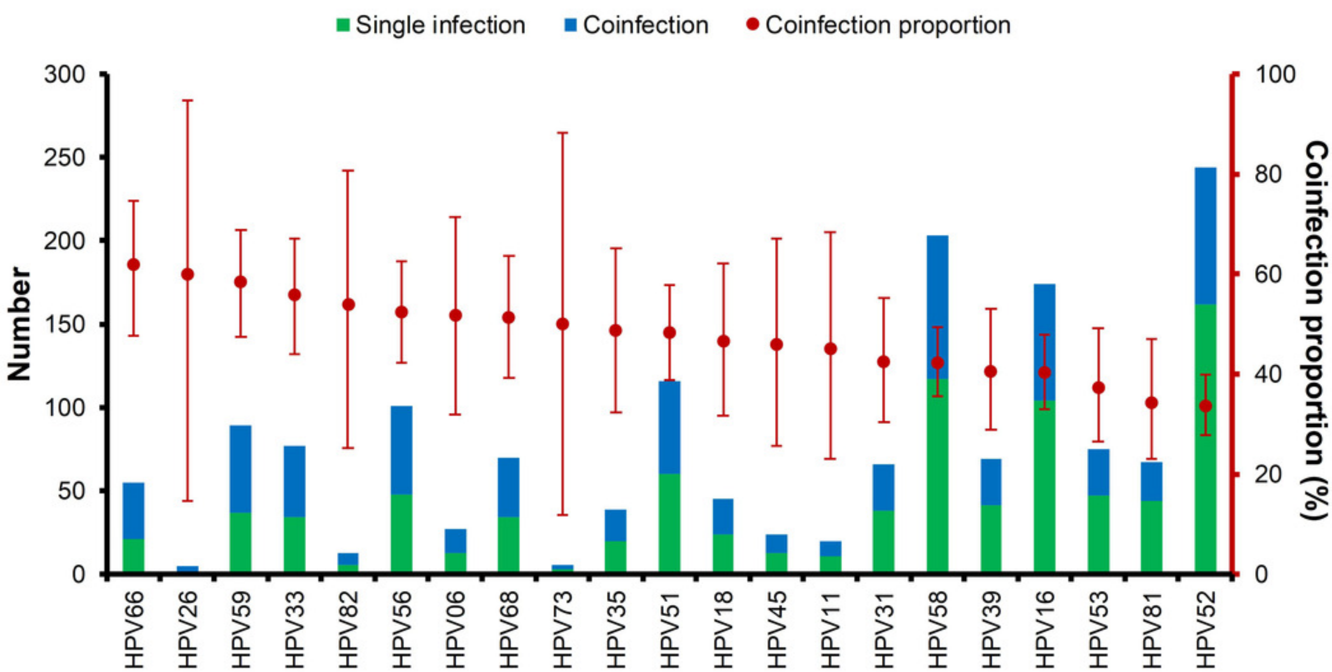

Table 1(on next page)

Table 1. The basic information of the subjects recruited at the different years in the study

Table 1. The basic information of the subjects recruited at the different years in the study

Table 1. The basic information of the subjects recruited at the different years in the study

Variable	Total	HPV (+)	HPV (-)	*P*
2014				
Age	43 (35-56)	46 (36-55)	43 (35-56)	0.380
Diabetes	105 (9.1)	13 (12.8)	92 (8.8)	0.181
Hypertension	206 (17.9)	22 (21.6)	184 (17.5)	0.307
2015				
Age	50 (38-64)	52 (39-63)	50 (38-64)	0.869
Diabetes	334 (14.4)	24 (14.5)	310 (14.4)	0.985
Hypertension	624 (26.9)	32 (19.3)	592 (27.51)	0.021
2016				
Age	45 (36-58)	45 (36-56)	45 (36-58)	0.351
Diabetes	431 (11.9)	30 (10.9)	401 (12.0)	0.594
Hypertension	795 (21.9)	59 (21.4)	736 (21.9)	0.831
2017				
Age	46 (36-60)	45 (35-56)	46 (36-60)	0.092
Diabetes	336 (10.9)	22 (8.6)	314 (11.2)	0.216
Hypertension	736 (24.0)	44 (17.3)	692 (24.6)	0.009

2018				
Age	44 (36-57)	45 (37-58)	43 (36-57)	0.330
Diabetes	466 (13.1)	36 (13.1)	430 (13.1)	0.997
Hypertension	775 (21.8)	60 (21.9)	715 (21.8)	0.980

3

Table 2(on next page)

Table 2. The HPV positive rates in the follow-up for different periods from the first HPV positive

Table 2. The HPV positive rates in the follow-up for different periods from the first HPV positive

Table 2. The HPV positive rates in the follow-up for different periods from the first HPV positive

Persistent period	Positive number	Total number	Positive rate (95% CI)	P
One year	186	218	85.3 (79.9-89.7)	<0.001
Two years	73	118	61.9 (52.5-70.6)	
Three years	73	118	61.9 (52.5-70.6)	

Note: The persist rate of HPV infection was calculated as (①+ and ②+)/(①+ and ②) for one year, (①+ and ②+ and ③+)/(①+ and ② and ③) for two years and (①+ and ②+ and ③+ and ④+)/(①+ and ② and ③ and ④) for three years, respectively, where ①, ②, ③, and ④ meant the detection at the first, second, third and fourth year, respectively.