Objective: To explore the genotype distribution of high-risk human papillomavirus (HR-HPV) and its attribution to different grades of cervical lesions in rural China, which will contribute to type-specific HPV screening tests and the development of new polyvalent HPV vaccines among the Chinese population.

Methods: One thousand two hundred ninety-two subjects were followed based on the Shanxi Province Cervical Cancer Screening Study I (SPOCCS-I), and screened by HPV DNA testing (hybrid capture\(^\text{\textregistered}\) 2 [HC2]), liquid-based cytology (LBC), and if necessary, directed or random colposcopy-guided quadrant biopsies. HPV genotyping with linear inverse probe hybridization (SPF10-PCR-LiPA) was performed in HC2 positive specimens. Attribution of specific HR-HPV type to different grades of cervical lesions was estimated using a fractional contribution approach.

Results: After excluding incomplete data, 1,274 women were included in the final statistical analysis. Fifteen point two percent (194/1,274) of women were HR-HPV positive for any of 13 HR-HPV types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68) and the most common HR-HPV types were HPV16 (19.1%) and HPV52 (16.5%). The genotypes most frequently detected in HR-HPV-positive cervical intraepithelial neoplasia grade 1 (CIN1) were HPV52 (24.1%), HPV31 (20.7%), HPV16 (13.8%), HPV33 (13.8%), HPV39 (10.3%), and HPV56 (10.3%); in HR-HPV-positive cervical intraepithelial neoplasia grade 2 or worse (CIN2+): HPV16 (53.1%), HPV58 (15.6%), HPV33 (12.5%), HPV51 (9.4%), and HPV52 (6.3%). HPV52, 31, 16, 33, 39, and 56 together contributed to 89.7% of HR-HPV-positive CIN1, and HPV16, 33, 58, 51, and 52 together contributed to 87.5% of CIN2+.

Conclusion: In summary, we found substantial differences in prevalence and attribution of CINs between different oncogenic HPV types in a rural Chinese population, especially for HPV16, 31, 33, 52, and 58. These differences may be relevant for both clinical management and the design of preventive strategies.

Keywords: Human Papillomavirus; Cervical Cancer; Cervical Intraepithelial Neoplasia; Genotype
INTRODUCTION

Cervical cancer is the fourth most common cancer in women worldwide, with an estimated 528,000 new cases and 266,000 deaths in 2012 [1]. A large majority (around 85%) of the global burden of cervical cancer occurs in less developed regions [1]. In 2015, there were an estimated 98,900 new cases and 30,500 deaths in China, accounting for respectively 18.7% and 11.5% of all cervical cancer cases and deaths worldwide [2]. It has been estimated that the annual number of new cases of cervical cancer may increase by approximately 40%–50% from 2010–2050 if no interventions are implemented based on the available epidemiological evidence from urban and rural areas in mainland China [3].

Persistent infection with high-risk human papillomavirus (HR-HPV) is the necessary cause of the cervical cancer and its precursors [4,5]. HR-HPV genotypes include HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68. HPV16 and 18 are the most common types worldwide [6,7]. World Health Organization (WHO) recommends HPV-DNA testing as a cervical cancer screening method if health resources allow in the area [8,9] as HPV infection is predictive of subsequent risk for developing cervical intraepithelial neoplasia [10,11]. Genotyping for HR-HPV may identify women at the greatest risk for developing cervical intraepithelial neoplasia grade 3 or worse (CIN3+) and may also warrant a less aggressive management of low-risk HPV infections [12].

There is considerable heterogeneity in the risk and distribution of HR-HPV types across regions worldwide, which may lead to differing cervical cancer incidences and mortalities among the countries [13,14]. Thus understanding the type-specific distribution of HR-HPV and its attribution to different grades of cervical lesions will direct the implementation of successful programs for cervical cancer prevention and management.

Xiangyuan County in Shanxi Province, China, is a high-risk area for cervical cancer, and has a high mortality rate. The first cervical cancer screening cohort, Shanxi Province Cervical Cancer Screening Study I (SPOCCS-I), was conducted in Xiangyuan in 1999. HR-HPV distribution and attribution to different grades of cervical lesions were evaluated in a 15-year follow up assessment.

MATERIALS AND METHODS

1. Participants

In 1999, 1,997 women aged 35–45 were enrolled in SPOCCS-I under the inclusion criteria as following: they were non-pregnant, had no history of cervical screening or hysterectomy, and were living in Xiangyuan County, Shanxi province. Each woman received HPV testing with hybrid capture® 2 (HC2), liquid-based cytology (LBC), visual inspection with acetic acid (VIA), colposcopy, 4-quadrant biopsy, and endocervical curettage (ECC). Women with histologically confirmed cervical intraepithelial neoplasia grade 2 or worse (CIN2+) lesions were offered immediate and affordable standard therapy on a voluntary basis. Treatment consisted of loop electrosurgical excision procedure, cone biopsy treatment, or hysterectomy, according to local guidelines. Details on participant recruitment and study design have been published previously [15,16]. The morbidity and mortality of cervical cancer among these women were routinely collected through cancer registry centers. Three additional follow-up visits were conducted in 2005, 2010, and 2014. All women who did not receive a hysterectomy
during 1999 to 2014 were invited to participate in the 2014 follow-up. Institutional Review Board approval was obtained from the Cancer Institute/Hospital of the Chinese Academy of Medical Sciences (CICAMS).

2. Screening procedure
In total, 1,292 eligible women signed informed consent forms and agreed to participate in the study in 2014. Information about cervical treatment history as well as sociodemographic characteristics and risk factors were collected via questionnaires. All samples of cervical exfoliated cells for women were collected by local health providers and transported to CICAMS for testing, which included LBC, HC2 testing, and HPV genotyping by experienced cytologists and senior technicians at CICAMS.

1) LBC
All patient samples underwent LBC (ThinPrep®, Hologic, Bedford, MA, USA) by experienced cytologists in CICAMS, and results were graded according to the Bethesda System. The cytological classifications were the following: negative for intraepithelial lesion or malignancy (NILM), atypical squamous cells of undetermined significance (ASC-US), atypical glandular cells (AGC), low-grade squamous intraepithelial lesion (LSIL), high-grade squamous intraepithelial lesion (HSIL), atypical squamous cells-cannot exclude HSIL (ASC-H), squamous cell carcinoma (SCC), adenocarcinoma in situ (AIS), and adenocarcinoma (ADC).

2) HC2
HC2 was performed at CICAMS by senior technicians according to the manufacturer’s instructions. RNA probes hybridized to single-stranded DNA were detected through antibody capture and chemiluminescence signal amplification. A positive result indicated the detection of any of the 13 types of HR-HPV (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68).

3) Linear inverse probe hybridization (SPF10-PCR-LiPA)
The automatic analyzer LiPA (INNO-LiPA™, Innogenetics, N.V., Ghent, Belgium) detected the HPV type(s) of an HC2-positive specimen using a reverse hybridization assay. This technology can detect 28 HPV types, including 13 types that are carcinogenic and/or identified as high-risk (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68), 7 types that are possibly carcinogenic (26, 53, 66, 69, 70, 73, and 82) and 8 types that are low-risk (6, 11, 40, 43, 44, 54, 71, and 74).

4) Colposcopy and pathological diagnosis
Women with cytological abnormalities (ASC-US or worse, ≥ASC-US) or HC2-positive results were called back for colposcopy visit by local experienced health providers. Biopsies were taken during this second follow-up visit as follows: directed biopsy was used in the case of colposcopically visible lesions; 4-quadrant punch biopsy was used for invisible lesions when cytological results were AGC, LSIL, HSIL, ASC-H, SCC, AIS, and ADC. ECC was performed if the squamocolumnar junction (SCJ) was not visualized. Histological diagnoses were categorized as the following: negative, cervical intraepithelial neoplasia grade 1 (CIN1), grade 2 (CIN2), grade 3 (CIN3), SCC, AIS, or ADC.

After excluding 18 cases which lacked colposcopy and biopsy results (2 cases with HR-HPV positive and ASC-US, 2 cases with LSIL, 13 cases with unsatisfactory cytology results, and...
1 case with unsatisfactory pathology result), 1,274 women were included in the final statistical analysis (Fig. 1).

3. Verification of disease status
Pathological diagnosis was used as the golden standard of diagnosis and CIN2+ was used as the clinical outcome endpoint. When pathological diagnosis was not available, a combination of secondary test results was used to verify the final disease status. According to previous studies on risk assessment in a Chinese population [8,15,16], women without biopsy were classified as negative if they fell into these categories: 1) negative colposcopy impression, 2) HR-HPV negativity with cytology negativity or ASC-US, or 3) HR-HPV positivity with cytology negativity; women without biopsy were classified as CIN2 if they fell into one of the following categories: 1) HSIL, or 2) ASC-H; women without biopsy were classified as “incomplete data” if they fell into one of the following categories: 1) LSIL, 2) HR-HPV positivity with ASC-US, or 3) AGC.

4. Statistical analysis
Socio-demographic characteristics of participants were described by using means, standard deviations (SD), frequencies, and proportions. Standard $\chi^2$ tests were conducted to compare...
the characteristics of different grades of cervical lesions (normal, CIN1, and CIN2+). The proportions of lesions caused by specific HPV types were estimated using the methods previously described [7,17]. In brief, the attributable fraction calculations proportionally weighted an HPV type in a multi-type infected lesion by the frequency of its occurrence as a single-type infection in a lesion of the same grade. For example, the derivation of an apportionment for 2 CIN2 lesions positive for both HPV16 and 52 in a study is as follows: if there were 5 single-type HPV16-infected CIN2 lesions and 3 single-type HPV52-infected lesions which also appeared in that study, then $2 \times 5/(5+3)$ or 1.25 of these 2 multi-type infected lesions would be attributed to HPV16 and $2 \times 3/(5+3)$ or 0.75 attributed to HPV52. All analyses were conducted using SPSS 19.0 (IBM Co., Armonk, NY, USA). Statistical tests were 2-sided, and p values less than 0.05 were considered statistically significant.

RESULTS

1. General characteristics
A total of 1,274 women were included in the final analysis. Mean age of participants was 54.5±3.6 years, mean age at menarche and first sexual intercourse were 16.0±1.8 and 21.0±2.2 years, respectively. 62.0% (789/1,274) had middle school or higher education, 10.1% (129/1,274) had more than 2 sexual partners, and 71.5% (911/1,274) were post-menopause. 4.9% (62/1,274) of women reported smoking, and 0.9% (12/1,274), reported drinking.

Among the 1,274 women, 1,199 (94.1%) were pathologically normal, 40 (3.1%) had CIN1, and 35 (2.7%) had CIN2+. From professional aspects and the characteristics distribution among the population, we set the following time points to be the criteria of which we classified participants into subgroups: 50 years old for age, 15 years old as an early age for menarche, and 20 years old as the early age for first sexual intercourse. No significant correlations were observed between demographic characteristics and cervical lesion severity (Table 1).

Table 1. Demographic characteristics of study population correlated to cervical lesion severity

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Normal</th>
<th>CIN1</th>
<th>CIN2+</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, yr</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>709 (59.1)</td>
<td>26 (65.0)</td>
<td>18 (51.4)</td>
<td>0.489</td>
</tr>
<tr>
<td>&gt;50</td>
<td>490 (40.9)</td>
<td>14 (35.0)</td>
<td>28 (48.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.450</td>
</tr>
<tr>
<td>Below middle school</td>
<td>451 (37.7)</td>
<td>19 (47.5)</td>
<td>13 (37.1)</td>
<td></td>
</tr>
<tr>
<td>Middle school or above</td>
<td>746 (62.3)</td>
<td>21 (52.5)</td>
<td>22 (62.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Age at menarche</strong>, yr</td>
<td></td>
<td></td>
<td></td>
<td>0.973</td>
</tr>
<tr>
<td>≤15</td>
<td>463 (38.7)</td>
<td>15 (37.5)</td>
<td>13 (37.1)</td>
<td></td>
</tr>
<tr>
<td>&gt;15</td>
<td>734 (61.3)</td>
<td>25 (62.5)</td>
<td>22 (62.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Age at first sexual intercourse</strong>, yr</td>
<td></td>
<td></td>
<td></td>
<td>0.126</td>
</tr>
<tr>
<td>≤20</td>
<td>553 (46.2)</td>
<td>25 (62.5)</td>
<td>16 (45.7)</td>
<td></td>
</tr>
<tr>
<td>&gt;20</td>
<td>644 (53.8)</td>
<td>15 (37.5)</td>
<td>19 (54.3)</td>
<td></td>
</tr>
<tr>
<td><strong>No. of sexual partners</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.218</td>
</tr>
<tr>
<td>≤1</td>
<td>1,078 (89.9)</td>
<td>38 (95.0)</td>
<td>29 (82.9)</td>
<td></td>
</tr>
<tr>
<td>&gt;1</td>
<td>121 (10.1)</td>
<td>2 (5.0)</td>
<td>6 (17.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Menopause status</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.989</td>
</tr>
<tr>
<td>Post-menopause</td>
<td>859 (71.6)</td>
<td>29 (72.5)</td>
<td>23 (65.7)</td>
<td></td>
</tr>
<tr>
<td>Pre-menopause</td>
<td>336 (28.0)</td>
<td>11 (27.5)</td>
<td>12 (34.3)</td>
<td></td>
</tr>
</tbody>
</table>

Values are presented as number (%).

CIN1, cervical intraepithelial neoplasia grade 1; CIN2+, cervical intraepithelial neoplasia grade 2 or worse.

*Two missing data; †Four missing data.
2. Distribution of HR-HPV among different grades of cervical lesions

In 194 women (15.2%, 194/1,274) with HR-HPV positive results, HPV16 (19.1%), 52 (16.5%), 58 (16.5%), 33 (13.4%), and 56 (11.9%) were the most common genotypes. The prevalence of HR-HPV in normal pathology, CIN1, and CIN2+ were 11.1% (133/1,199), 72.5% (29/40) and 91.4% (32/35), respectively. HR-HPV prevalence is positively correlated with cervical lesion severity ($\chi^2$=286.979; p<0.001).

Among pathologically normal women, the 5 most prevalent HR-HPV types were HPV58 (19.6%), HPV52 (17.3%), HPV56 (14.3%), HPV51 (14.3%), and HPV33 (13.5%); among CIN1: HPV52 (24.1%), HPV31 (20.7%), HPV16 (13.8%), HPV33 (13.8%), HPV39 (10.3%), and HPV56 (10.3%); among CIN2+: HPV16 (53.1%), HPV58 (15.6%), HPV33 (12.5%), HPV51 (9.4%), and HPV52 (6.3%).

The proportions of single and multiple infections in women were as follows: 77.4% (103/133) and 22.6% (30/133) in normal pathology, 82.8% (24/29) and 17.2% (5/29) in CIN1, 87.5% (28/32) and 12.5% (4/32) in CIN2+. The trend test showed no significance. HPV16 was the most common genotype among single infections in CIN2+ (Fig. 2).

The prevalence of HPV16 was positively correlated with the severity of cervical lesions (12.0% in normal, 13.8% in CIN1, 53.1% in CIN2+, $\chi^2$=28.838; p<0.001). No significant correlation between HR-HPV prevalence and CIN1/CIN2+ diagnosis was observed except in HPV16 ($\chi^2$=5.063; p=0.024) and HPV31 (Fisher, p=0.046).

3. The attributable proportion of HR-HPV in different grades of cervical lesions

Table 2 presents the prevalence of specific HR-HPV types in different grades of HR-HPV-positive cervical lesions along with an estimate of the attributable fraction of HR-HPV (i.e., an estimate of the proportion of lesions caused by a given HPV type). A total of 11 and 10 HR-HPV types were detected in CIN1 and CIN2+ respectively.

HPV16 was attributed to 13.8% of CIN1 and 52.2% of CIN2+, and was ultimately the most prevalent genotype observed in CIN2+. In total, HPV52, 31, 16, 33, and 56 combined

Table 2. Distribution and attributable proportion of HR-HPV genotypes in different grades of cervical lesions

<table>
<thead>
<tr>
<th>HR-HPV genotype</th>
<th>Total population (n=1,274)</th>
<th>Normal pathology (n=1,199)</th>
<th>CIN1 (n=40)</th>
<th>CIN2+ (n=35)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>Prevalence (%)</td>
<td>Single infection (%)</td>
<td>Attributable proportion (%)</td>
</tr>
<tr>
<td>Any type</td>
<td>194</td>
<td>100.00</td>
<td>133</td>
<td>100.00</td>
</tr>
<tr>
<td>16</td>
<td>37</td>
<td>19.07</td>
<td>27</td>
<td>16.70</td>
</tr>
<tr>
<td>18</td>
<td>14</td>
<td>7.22</td>
<td>8</td>
<td>5.53</td>
</tr>
<tr>
<td>31</td>
<td>18</td>
<td>9.28</td>
<td>10</td>
<td>6.83</td>
</tr>
<tr>
<td>33</td>
<td>26</td>
<td>13.40</td>
<td>19</td>
<td>11.79</td>
</tr>
<tr>
<td>35</td>
<td>5</td>
<td>2.58</td>
<td>1</td>
<td>1.10</td>
</tr>
<tr>
<td>39</td>
<td>13</td>
<td>6.70</td>
<td>11</td>
<td>6.43</td>
</tr>
<tr>
<td>45</td>
<td>5</td>
<td>2.58</td>
<td>4</td>
<td>2.10</td>
</tr>
<tr>
<td>51</td>
<td>22</td>
<td>11.34</td>
<td>11</td>
<td>8.02</td>
</tr>
<tr>
<td>52</td>
<td>32</td>
<td>16.49</td>
<td>24</td>
<td>15.26</td>
</tr>
<tr>
<td>56</td>
<td>23</td>
<td>11.86</td>
<td>13</td>
<td>9.05</td>
</tr>
<tr>
<td>58</td>
<td>32</td>
<td>16.49</td>
<td>20</td>
<td>13.18</td>
</tr>
<tr>
<td>59</td>
<td>6</td>
<td>3.09</td>
<td>5</td>
<td>2.58</td>
</tr>
<tr>
<td>68</td>
<td>6</td>
<td>3.09</td>
<td>2</td>
<td>1.44</td>
</tr>
</tbody>
</table>

CIN1, cervical intraepithelial neoplasia grade 1; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; HPV, human papillomavirus; HR-HPV, high-risk human papillomavirus.

*The sum of the percentages of each HPV type is not necessarily equal to 100% because a result may be counted more than once in cases where the sampled lesion contained multiple HPV types; †Proportion of multiple attributable fraction and number of single infection within HR-HPV positive women.
Fig. 2. The most common HR-HPV genotypes in different grades of cervical lesions. The sum of the percentages of each HPV type in 1 type of cervical lesions is not equal to 100% because a result may be counted more than once if the lesion contained multiple HPV types.

CIN1, cervical intraepithelial neoplasia grade 1; CIN2+, cervical intraepithelial neoplasia grade 2 or worse; HPV, human papillomavirus; HR-HPV, high-risk human papillomavirus.
were attributed to 89.7% of all CIN1 lesions, and HPV16, 33, 51, 52, and 58 combined were attributed to 87.5% of all CIN2+ lesions.

All HPV types in different lesions had lower attribution than their prevalence after weighting an HPV type in a co-infected lesion. While the ranking of HR-HPV types between prevalence and attributable proportion differed among lesions, the top 5 HR-HPV types remained constant.

**DISCUSSION**

In this cross-sectional population-based study of 1,274 women in Shanxi province, the overall HR-HPV prevalence was 15.2%. This rate is consistent with that previously reported in China (17% in Shenyang [18], 18% in Shenzhen [19], and 19% in Jiangsu [20]). The most frequent HPV genotypes identified in our study population were 16, 52, and 58. These results are consistent with most of other studies conducted in China which routinely found HPV16 to be the most common genotype in the general population [21,22]. However, HPV52 was the most commonly detected HR-HPV types in cervical specimens in Jiangsu, Guangzhou, and some other provinces [18,23]. Furthermore, some areas in Asia such as India have a similar HR-HPV prevalence to our results, with HPV16, 18, and 58 being the most common HPV types there [24]. A study exploring the distribution of HPV genotypes over 5 continents and 59 countries found that the main HPV genotypes in women with normal cytological results were 16, 18, 52, and 31 [25]. A worldwide meta-analysis of over 115,000 HPV-positive women also found that while HPV16 and 18 were the most prevalent types of HPV associated with cervical lesions worldwide, HPV52 and 58 were more prevalent in East Asian women [7].

Quantifying the proportion of cervical lesions caused by each HPV type is a critical parameter for exploring the HPV-specific risks of developing high-grade CIN, SCC, or cervical ADC [13,14,26,27]. In the present study, HR-HPV DNA prevalence increased with the severity of pathological abnormalities. HPV16 was the predominant HR-HPV genotype and was attributable to 52.2% of CIN2+ lesions. Other important genotypes correlating to CIN2+ lesions were HPV33, 51, 52, and 58. HPV52, 31, 16, 33, 39, and 56 were most prevalent in CIN1, totaling 89.7% of all CIN1 lesions. These results are consistent with those of Lu et al. in Beijing [28] and the study on Chinese women in 2013 which showed that 71.4% of HR-HPV positive CIN2+ cases were attributable to HPV16 or 18 [29]. A 14-year follow-up study in Sweden presented HPV16 and 31 as contributing to higher incidences of CIN2+ [11,30].

Phase III clinical trials for 2 widely used prophylactic HPV vaccines (HPV16/18 and HPV16/18/6/11) have been ongoing in mainland China for over 7 years, and the bivalent vaccine has been recently approved in China. Therefore, it is important to determine the distribution of HPV types in different cervical lesions in order to estimate the potential protection provided by current HPV vaccines. In addition, the determination of the most common HPV types in cervical lesions can influence the development of second-generation polyvalent HPV vaccines. Our study suggests that HPV16, 31, 33, 52, and 58 play important roles in the development of the cervical lesions in rural China, and these particular HPV carcinogenic types should be given priority when assessing the cross-protective effects of HPV vaccines and developing type-specific high-risk HPV DNA-based screening tests and strategies in rural China.

HPV18 has an established association with ADC [31], but it is more frequently present in cancer rather than in precancerous lesions [32]. As a result, its prevalence and attribution
to the cervical lesions may have been under-represented in our cohort due to the limited number of cervical cancer incidences, and the low prevalence in our study may have limited our ability to reach a conclusion on HPV18's attributable risk in this population.

In conclusion, given the differing risks attributed to HPV type, we propose that healthcare workers should pay attention to the specific HPV types in cervical lesions during screening in the rural Chinese population, especially for HPV16, 31, 33, 52, and 58, as this may convey valuable information regarding cervical cancer risk. Furthermore, the findings will help inform clinicians and health providers about appropriate personalized management of screening abnormalities and follow-up strategies for the screening population.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr Youlin Qiao, Wenhua Zhang, Qinjing Pan, Xun Zhang, Jerome L Belinson, and other investigators from SPOCCS-I for the enrollment study in 1999; and all the women participants and local workers contributed to the follow-ups in the year of 2014.

REFERENCES


