BMJ Open

Estimating the prevalence of non-cervical human papillomavirus infection in mainland China (PROGRESS-Plus): protocol of a national cross-sectional study

Wei Wang ,[1] Smita Kothari-Talwar,[1] Craig Roberts,[1] Qingbin Lu,[2] Nuria Lara,[3] Ravi Panchal,[4] Edith Morais,[5] Ya-Ting Chen,[1] Youlin Qiao,[6] Fuqiang Cui [7]

To cite: Wang W, Kothari-Talwar S, Roberts C, *et al*. Estimating the prevalence of non-cervical human papillomavirus infection in mainland China (PROGRESS-Plus): protocol of a national cross-sectional study. *BMJ Open* 2023;**13**:e065749. doi:10.1136/bmjopen-2022-065749

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-065749).

Received 16 June 2022
Accepted 17 January 2023

[1]Merck & Co Inc, Rahway, New Jersey, USA
[2]School of Public Health, Peking University, Beijing, China
[3]IMS Health Barcelona, Barcelona, Spain
[4]IQVIA, Bangalore, India
[5]MSD France, Courbevoie, France
[6]Cancer Epidemiology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[7]Peking University Health Science Centre, Beijing, China

Correspondence to
Dr Wei Wang;
wei.wang40@merck.com

ABSTRACT

Introduction Human papillomavirus (HPV) infection is the most common sexually transmissible infection worldwide. Although the prevalence of cervical HPV infection has been extensively reported in women worldwide, few epidemiological studies have examined the prevalence of non-cervical HPV infection among both women and men, especially in China.

Methods and analysis PROGRESS-Plus is a national, multisite, cross-sectional study that aims to estimate the prevalence of non-cervical HPV infection in women and men aged 18–60 years residing in mainland China. More specifically, PROGRESS-Plus will estimate the prevalence rate of HPV DNA in oral samples from both women and men, and that of anogenital samples from men. The secondary study objectives are to (1) report the aforementioned prevalence rates by HPV genotype, age and geographical region, (2) examine the concordance (ie, prevalence of the same HPV genotype) between the oral and anogenital samples among men, (3) explore risk factors associated with oral (in both women and men) and anogenital (in men only) HPV infection and (4) describe study participants' health-related quality of life, health behaviour, sexual behaviour and health status.

Ethics and dissemination The study protocol and all required documents have been submitted for review and approval to the Independent Ethics Committees of all the participating sites. All participants will provide their written informed consent on study entry, and all the recorded data will be treated as confidential.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Collaborating with local Centers for Disease Control and Prevention (CDC) and labs, PROGRESS-Plus is a nationally representative, multisite, cross-sectional study of human papillomavirus prevalence in the general Chinese population.
⇒ To obtain a representative sample of a geographically diverse population, PROGRESS-Plus uses a two-step sampling approach that combines probability proportional to size and targeted community sampling, which maximises the generalisability of study results across the country.
⇒ The standardised processes and approaches established in PROGRESS-Plus will serve as a methodological model or framework for future epidemiological and public health research on a variety of health outcomes in mainland China.
⇒ PROGRESS-Plus will collect data on patient-reported sexual behaviour, where there will be potential missing responses in the participant surveys.

BACKGROUND

Human papillomavirus (HPV) infection is the most common sexually transmissible infection worldwide.[1 2] Although most HPV infections cause no symptoms and resolve on their own in 1–2 years, some infections are persistent and can lead to mucosal and skin lesions, such as genital warts and cancer.[3 4] A prominent example of HPV-related diseases is cervical cancer, the fourth most frequently diagnosed cancer and the fourth leading cause of cancer death among women worldwide.[5] Over 90% of the global cervical cancer cases are caused by HPV.[6] HPV 16 and 18, two of the high-risk HPV (HR-HPV) genotypes, alone are responsible for over 70% of the cervical cancers globally. Furthermore, the oncogenic effects of HPV have been observed in many anatomic sites other than the cervix. It is estimated that non-cervical HPV infection is responsible for 50% of penile, 88% of anal, 43% of vulvar, 70% of vaginal and 13%–56% of oropharyngeal cancers.[7 8]

There is a large body of literature on the prevalence of HPV infection in various populations and geographical locations. In a recent meta-analysis of 66 studies, the prevalence of oral HPV infection (any genotype) was estimated to be 7.7% (95% CI 6.8% to 8.6%) worldwide, with significantly higher prevalence being observed among men (9.3%; 95% CI 6.4% to 12.5%) compared with women (5.5%; 95% CI 4.5% to 6.6%).[9] The highest rate of oral HPV infection was

observed in South America (12.4%; 95% CI 7.2% to 12.5%) and followed by Europe (9.9%; 95% CI 4.5% to 6.6%), whereas the lowest rate was observed in Asia (2.6%; 95% CI 0.6% to 4.6%).[9] Regarding the prevalence of oral infection with HR-HPV, a recent meta-analysis of 48 studies estimated that the prevalence was 4.9% (95% CI 3.7% to 6.1%) worldwide, with no significant sex differences.[10] On the other hand, the prevalence of genital HPV infection is higher than that of oral HPV infection. Among women with normal cytological findings, the global prevalence of genital HPV infection was 11.7% (95% CI 11.6% to 11.7%) according to a meta-analysis of 194 studies.[11] This was consistent with the findings from another meta-analysis which estimated that the HPV prevalence of genital HPV infection was 11% (95% CI 9% to 12%) in asymptomatic women based on 32 studies.[12] Regarding geographical differences, the prevalence of genital HPV infection in women was the highest in sub-Saharan Africa (24.0%; 95% CI 23.1% to 25.0%) and Eastern Europe (21.4%; 95% CI 20.1% to 22.7%).[11] Among men, the prevalence of genital HPV infection was 12.4% (95% CI 5.6% to 21.5%) in the European population according to a meta-analysis of 31 studies.[13] In the USA, it was estimated that the overall prevalence of genital HPV infection among men was 45.2% (95% CI 41.3% to 49.3%).[14]

In mainland China, there is no official HPV vaccine policy or guideline, and the decision to get vaccinated for HPV was made by consumers. The prevalence of cervical HPV infection has been extensively reported in women from many geographical regions in China.[15] However, there is a lack of epidemiological studies that examine the prevalence of non-cervical HPV infection among both women and men. For instance, to our knowledge, only one single-site study reported the rate of oral HPV infection in mainland China.[16 17] In this study of 5410 healthy adults from rural Anyang city, Hang *et al* found that the prevalence of α mucosal HPV and oncogenic α mucosal HPV was 0.67% (95% CI 0.47% to 0.93%) and 0.50% (95% CI not reported), respectively, with no significant sex differences.[16] In addition, only two single-site studies reported the rate of anogenital (other than cervical) HPV infection in mainland China.[18–21] In the first study,[18–20] Wei *et al* recruited 4687 healthy adults from Liuzhou city in Guangxi, and collected specimens from the vulva, vagina and perianal/anal canal (PA) in women and penis/glans penis/coronary sulcus (PGC) and PA in men. The investigators found that the overall prevalence of HPV infection in these anatomic sites was 15.4% (95% CI 14.4% to 16.6%).[18] Compared with men, women had significantly higher prevalence of overall (19.5% in women vs 10.5% in men) and oncogenic HPV infection (18.7% in women vs 9.4% in men) in these anatomic sites.[18] In the second study,[21] He *et al* enrolled 3172 men from rural Anyang in Henan and collected specimens from PGC and scrotum. The investigators found that the prevalence of overall HPV and oncogenic HPV infection was 17.5% (95% CI 16% to 19%) and 6.3% (95% CI 5.3% to 7.3%), respectively, in these anatomic sites in the study samples.[21] A few other studies have examined the prevalence of HPV infection in healthy individuals in China, however, they focus on specific population subgroups rather than the general Chinese population, such as men who have sex with men.[22 23] Overall, there exists limited evidence on non-cervical (oral and anogenital) HPV infection in mainland China.

To fill the knowledge gap in non-cervical HPV infection in mainland China, we propose to conduct the prevalence of oral HPV infection, a global assessment, plus the prevalence of anogenital HPV infection among men (PROGRESS-Plus) study in China. The primary goal of PROGRESS-Plus is to examine the prevalence of oral HPV infection among men and women, and the prevalence of anogenital HPV infection among men in the general adult population in mainland China.

METHODS

Study design and objectives

PROGRESS-Plus is a 3-year national, multisite, cross-sectional study that is scheduled to begin in July 2021 and end in June 2024. The primary study objective is to estimate the prevalence of non-cervical HPV infection in women and men aged 18–60 years residing in mainland China (ie, target population; excluding Hong Kong, Macau and Taiwan). More specifically, PROGRESS-Plus will estimate the prevalence rates of HPV DNA in oral samples from both women and men, and anogenital samples from men. The secondary study objectives are to (1) report the aforementioned prevalence rates by HPV genotype, age and geographical region, (2) examine the concordance (ie, prevalence of the same HPV genotype) between the oral and anogenital samples among men, (3) explore risk factors associated with oral (in both women and men) and anogenital (in men only) HPV infection and (4) describe study participants' health-related quality of life (HRQOL), health behaviour, sexual behaviour and health status.

Sampling strategy and participant recruitment

A two-stage sampling approach that employs a combination of probability proportional to size (PPS) and targeted community sampling (TCS) will be used. First, following the PPS sampling method, 12 study sites will be drawn from urban (100 000 or more inhabitants) and rural areas (fewer than 100 000 inhabitants) in 6 regions in China (North, Northeast, East, Southcentral, Southwest, Northwest) (figure 1). Second, participants will be recruited into the study using TCS, which involves specific community outreach that invites those who are interested in the study to present themselves for eligibility screening at participating study sites. The targeted numbers of study participants have been established by region, type of habitat (urban or rural), age groups and sex (table 1), with the goal of producing a sample that reflects the age,

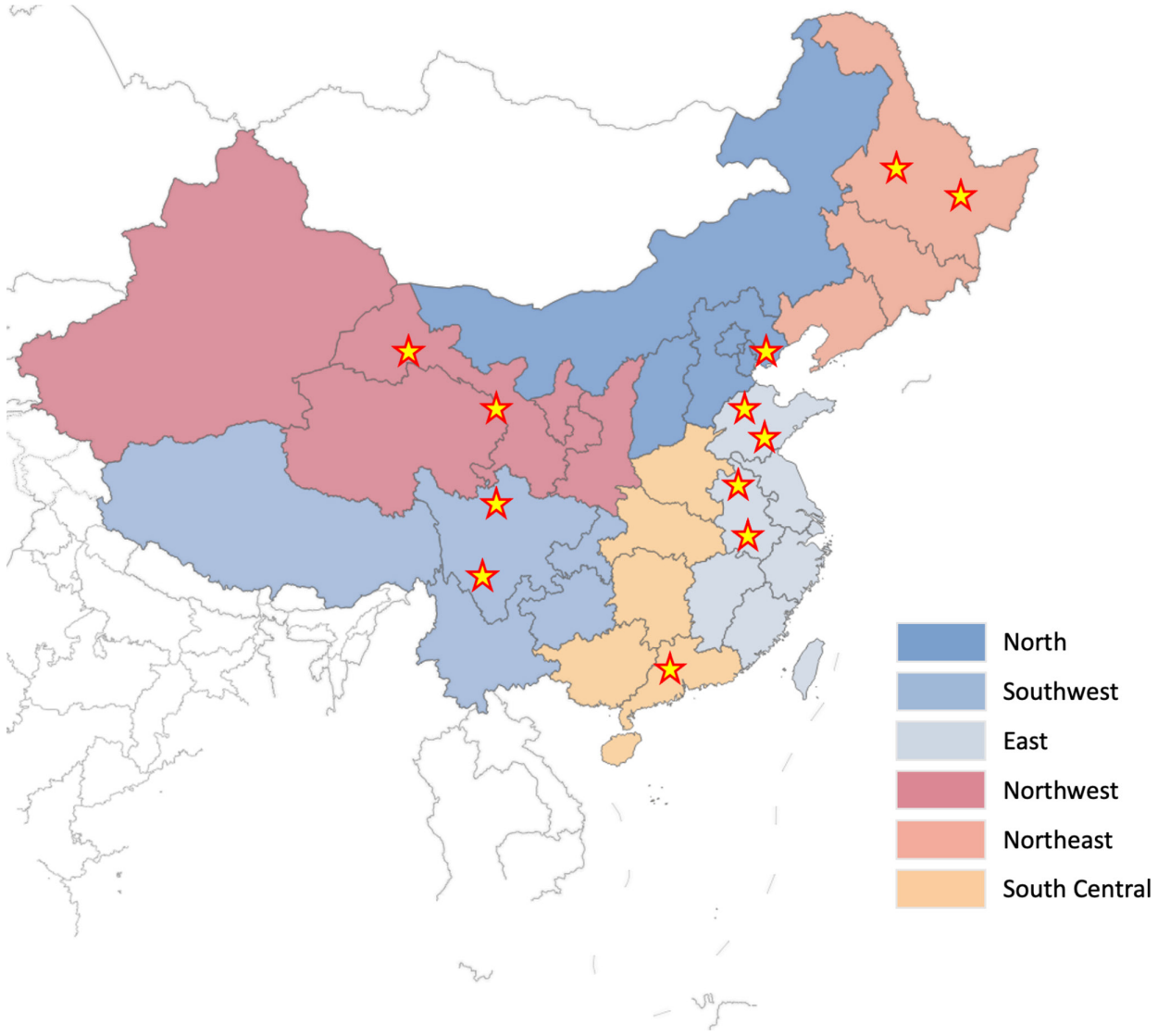

Figure 1 Participating study sites in PROGRESS Plus. Figure 1 was created using the Tableau software, which contains country shapefiles.

sex and rural versus urban distribution of the general population in China.

Sample size calculation

Sample size calculation was performed to determine the number of participants required to achieve a desirable precision level for the prevalence estimates at the 95% confidence level using formula (1).

$$n = 1.96^2 \times \frac{p(1-p)}{d^2} \quad (1)$$

In (1), p is the expected prevalence rate and d is the precision level. The HPV prevalence rates reported in the literature and varying precision levels were used as inputs to calculate the required sample sizes under different conditions. In addition, an estimated 20% biological sample loss was considered in the calculation. Potential reasons for sample loss included participants not accepting some or all specimen collections after consenting or collected specimens not having enough HPV DNA to provide conclusive results.

In the general Chinese population, the prevalence of oral HPV infection was estimated at 0.67% (95% CI 0.47% to 0.93%), 2.5% (95% CI 1.8% to 3.5%) and 6.7% (95% CI 5.6% to 7.7%) for all genotypes combined in prior studies.[16 24] These prevalence rates and higher upper 95% CI limit (7.7%) were entered in formula (1) to calculate sample sizes under precision levels 0.35%–2.2% depending on prevalence rate (ie, higher precision level was used for lower prevalence estimates). It was determined that a sample size of 4742 participants (2371 women and 2371 men) was desired as it allowed the estimation of prevalence rates 0.67%, 2.5%, 6.7% and 7.7% at precision levels 0.35%, 0.7%, 1.1% and 1.2%, respectively. The precision level for each estimation was selected to balance precision of the estimation and feasibility. The prevalence of anogenital HPV infection was estimated at 12.6%, 14.5% and 16.9% for all genotypes combined.[15 19 25] One study also reported that the prevalence of HPV infection was 10.8% in the PGC areas and 3.8% in the PA areas.[19] Based on formula

Table 1 Targeted numbers of study participants by region, site location, age and sex

Region	18–30 years old			31–45 years old			46–60 years old			
	Overall	Male	Female	Overall	Male	Female	Overall	Male	Female	Overall
North										
Rural	60	30	30	90	45	45	90	45	45	240
Urban	90	45	45	132	66	66	134	67	67	356
Northeast										
Rural	34	17	17	52	26	26	52	26	26	138
Urban	58	29	29	86	43	43	86	43	43	230
East										
Rural	124	62	62	186	93	93	188	94	94	498
Urban	224	112	112	334	167	167	338	169	169	896
Southcentral										
Rural	138	69	69	206	103	103	208	104	104	552
Urban	200	100	100	298	149	149	300	150	150	798
Southwest										
Rural	82	41	41	122	61	61	124	62	62	328
Urban	90	45	45	134	67	67	134	67	67	358
Northwest										
Rural	40	20	20	60	30	30	62	31	31	162
Urban	46	23	23	70	35	35	70	35	35	186
Total	1186	593	593	1770	885	885	1786	893	893	4742
Rural	478	239	239	716	358	358	724	362	362	1918
Urban	708	354	354	1054	527	527	1062	531	531	2824

(1), a total sample of 2371 men allowed the estimation of prevalence rates 10.8% for PGC, 3.8% for PA and 16.9% overall at precision levels 1.4%, 0.8% and 1.6%, respectively. Considering the above, it was planned that PROGRESS-Plus would enrol a minimum sample of 4742 participants (2371 women and 2371 men) from the target population.

Inclusion and exclusion criteria

To be eligible for PROGRESS-Plus, an individual must satisfy the following criteria: (1) between the ages of 18 and 60, (2) not actively seeking healthcare due to a sexually transmittable disease (STD) or lesions which may be caused by HPV as per the attending physician assessment, (3) being able to comprehend and answer the questions in the participant survey and (4) being able to provide written informed consent. To ensure the representativeness of PROGRESS-Plus study population as compared with the general Chinese population, individuals with known HPV-related diseases are not excluded from this study.

Patient and public involvement

The survey questionnaire has been pilot tested in a group of volunteers. Individuals will participate in the study after giving written informed consent.

Study procedures

Overview

When individuals present themselves at the participating study sites, they will receive an information sheet about the study and subsequently be assessed for eligibility by a physician. Eligible individuals who agree to participate and sign an informed consent will complete a survey questionnaire and provide a sample for HPV testing. Female participants will provide an oral rinse and gargle (ORG) sample, and male participants will provide one ORG and two anogenital swab samples. Participant recruitment is expected to last up to 52 consecutive working weeks, or until all sites have reached the targeted number of participants. The study protocol and all required documents have been submitted for review and approval to the independent ethics committees (IEC) of all the participating sites. In addition to IEC approval, regulatory authority approval (Office of Human Genetic Resource Administration) has been obtained in accordance with local regulatory requirements. All participants will provide their written informed consent on study entry, and all recorded data will be treated as confidential.

ORG sampling procedure

This study requires all women and men to provide an ORG sample during the study visit. Using an oral rinse,

participants will be instructed to do a mouth wash and gargle to collect cells from both the oral cavity and the oropharynx. The ORG sample will be collected in an Eppendorf tube and stored in the freezer at –20°C or lower until shipment to the laboratory for testing. The oral samples will be delivered on dry ice to a central lab for HPV PCR testing and storage at –80°C.

Anogenital sampling procedure

Only men will be required to provide anogenital samples. Dacron or flocked swabs moistened with sterile saline will be used to collect three specimens from the external genitalia (penis (coronal sulcus/glans penis/penis shaft), scrotum and perineal/perianal region), and one from anal canal by trained technicians. The penis (coronal sulcus/glans penis/penis shaft), scrotal skin and perineal/perianal skin will be gently abraded with a nail file and then swabbed. For uncircumcised men, exfoliated cells of the foreskin will be collected using the same swab used for penis sample. A fourth swab will be gently inserted into the anus as far as it will go until resistance is met (generally 5–6 cm), swabbing the inside of the anal canal in a 360° rotating motion 2–3 times. The four swab samples will be placed into four separate containers of Specimen Transport Medium (STM, DIGENE, Gaithersburg, Maryland, USA) and stored at –20°C (or lower) immediately after collection at the site until shipment. The anogenital samples will be delivered on dry ice to a central lab for HPV PCR testing and storage at –80°C. Details of sample collection and storage will be documented in laboratory manual or separate operation plans.

Central laboratory testing

HPV DNA detection and genotype identification will be performed using the SPF_{10} LiPA assay, a commercially available, highly sensitive test. To validate the presence of DNA and possible detection of PCR inhibition, RNase P DNA internal control validation will be performed. The DNA enzyme immuno-assay (DEIA) SPF_{10} method will be used to assess HPV DNA and genotype. This method permits highly sensitive, broad-spectrum detection of HPV DNA. The SPF_{10}-PCR products will be analysed by hybridisation to a cocktail of general HPV probes. This permits general detection of HPV-DNA in a DEIA. Genotyping will then take place by means of the INNO-LiPA HPV genotyping assay. This assay is based on the principle of reverse hybridisation.

Variables of interests

PROGRESS-Plus will collect three different types of data from the participants: (1) clinical data reported by the attending physicians, (2) self-reported data on sociodemographics, health behaviour, HRQOL, sexual behaviour, and disease status and history from participant surveys and (3) HPV DNA and genotypes data from laboratory tests on oral and anogenital biological samples. A complete list of variables is provided in table 2.

Clinical data reported by attending physicians

Participants' clinical data reported by the attending physicians will include information on the clinical site and attending physician, and participants' general health and HPV vaccination status. Clinical site information includes habitat (urban or rural), type of healthcare centre (county hospital, township hospital, community healthcare centre, public or private clinic, other) and type of site (private or public). Physician information includes clinical specialty and years in practice. Participants' general health information (eg, age and sex), along with participants' HPV vaccination status, including age at vaccination, type of vaccine received, number of doses and dates of administration, will be provided by participants and documented by physicians in the clinical report forms. Participants' general health of the oral cavity, number of natural missing teeth and presence of periodontal disease will be examined at the visit and documented by trained physicians.

Self-reported data from participant surveys

Sociodemographic data to be collected in the participant survey will include region and country of birth, education, marital status, employment status and health insurance status. Health behaviour data will include smoking habit and recent alcohol consumption. Survey questions on smoking habit will assess the type of tobacco product (cigarettes or other types), average number of cigarettes smoked per day, age of smoking initiation, number of years of smoking and number of years since quitting for former smokers (pack-years will be calculated). Questions on recent alcohol consumption will assess the type of alcohol product, and drinking frequency and quantity. Participants' HRQOL will be evaluated using the five-level EuroQol-5 Dimension instrument.[26] Sexual behaviour data will include the type of sexual behaviour (having sex with men, women or both), and lifetime and recent sexual activities. Survey questions on sexual activities will assess the age at first sexual activity and number of lifetime sexual partners (men and women). Participants will be asked to indicate whether they have done any oral, genital/vaginal and anal (receptive and/or insertive) sexual activities (lifetime and recent) in separate questions. For each type of sexual activity they indicate having done (eg, oral sex), participants will be asked about age at first such sexual activity, number of sex partners for such sexual activity, and last time of such sexual activity. Men who report having anogenital sex with other men will be asked to provide the number of sex partners who they had insertive and/or receptive anal sex. Lastly, disease status and history data will be collected. The survey questions on genital warts and cervical dysplasia will assess whether the participants ever had genital warts, ever had a partner with genital warts, ever had cervical dysplasia or abnormal cervical testing within the last 12 months (women only), and ever had a female partner with cervical dysplasia or abnormal Pap test. The survey questions on cancer history will assess oropharyngeal cancer in both men and women, penile and anal cancer in men, and vaginal, vulvar

Table 2 Variables to be collected in progress plus

Data source	Data domain	Variables
Clinical data reported by attending physicians	Site information	1. Habitat (urban or rural) 2. Type of healthcare centre (county hospital, township hospital, community healthcare centre, public or private clinic, other) 3. Type of site (private, public)
	Physician data	1. Clinical specialty (GP, dermatologist, urologist, etc) 2. Years attending patients at any site
	General health assessment	1. Age 2. Sex 3. General health of the oral cavity 4. No of natural missing teeth 5. Presence of periodontal disease
	HPV vaccination	1. HPV vaccination status (vaccinated, not vaccinated) 2. Type of vaccination 3. Dose of vaccination 4. Date of vaccination
Participant self-reported data	Socio-demographics	1. Region/country of birth 2. Years of school completed 3. Marital status 4. Health insurance
	Health-related quality of life	1. Mobility 2. Self-care 3. Usual activities 4. Pain/discomfort 5. Anxiety/depression
	Health behaviour	1. Smoking habits: cigarettes or other forms of tobacco such as cigars, and tobacco chewing 2. Alcohol consumption: recent alcohol consumption and quantity
	Sexual behaviour	1. Sexual behaviour: sex with men, sex with women, both. 2. Lifetime and recent sexual activity: if yes, age at first sexual activity; no of lifetime sexual partners, men and women; no of sex partners for insertive/receptive anal sex 3. Lifetime and recent oral sexual activity: if yes, age at first oral sex; number of oral sex partners, men and women; last time of any oral sex 4. Lifetime and recent genital/vaginal sexual activity: if yes, age at first genital/vaginal sex; no of genital/vaginal sex partners, men and women; last time of any genital/vaginal sex 5. Lifetime and recent anal (receptive and/or insertive) sexual activity: if yes, age at first anal sex; no of anal sex partners, men and women; last time of any anal sex
	Disease history	1. Genital warts and cervical dysplasia: ever had genital warts; ever had a partner with genital warts; for females: ever had cervical dysplasia/abnormal Pap test and Pap test within the last 12 months; ever had a female partner with cervical dysplasia or abnormal Pap test 2. Cancer history: oropharyngeal (men and women); penile and anal (men); vaginal, vulvar and cervical (women) 3. Other conditions: history of tonsillectomy; presence of warts in the mouth or throat; presence of anogenital warts and/or condyloma (men); medical history: concomitant diseases and sexually transmitted diseases, weakened immune system due to any disease or treatment
Laboratory test data	HPV DNA and genotypes	1. In the ORG sample 2. In the anogenital samples, for males only

GP, general practitioner; HPV, human papillomavirus; ORG, oral rinse and gargle.

and cervical cancer in women. Participants will also be asked about their history of tonsillectomy, presence of warts in the mouth or throat, presence of anogenital warts and/or condyloma (men only), history of concomitant diseases and STDs, and presence of a weakened immune system due to any disease or treatment (ie, HIV or AIDS, use of immune system-suppressing drugs)

HPV DNA and genotypes data from laboratory tests

The presence of HPV DNA and its genotype will be identified using the SPF_{10} LiPA assay. All participants will be classified as HPV positive or negative. HPV genotypes will be determined for HPV-positive participants. Genotypes 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59 are considered HR-HPV. Genotypes 6 and 11 are considered LR-HPV.

Data analysis

Primary analysis

All data analysis will be conducted using the newest version of SAS 9.4 or R. To characterise the study sample, descriptive statistics will be reported for all variables of

interest, including frequencies and percentages for categorical variables, and mean, median, SD and IQR for continuous variables. To calculate the prevalence of oral HPV infection, the number of participants with an oral HPV infection will be divided by the total number of study participants with an available ORG sample. To calculate the prevalence of anogenital HPV infection among men, the number of participants with HPV infection in each anogenital area (external genitalia and anal canal) will be divided by the total number of study participants with an available sample from the corresponding area. All prevalence rates will be reported with a 95% CI.

Secondary analysis

First, the prevalence of oral HPV infection will be reported by HPV genotype, sex (women, men), age groups (18–24 years, 25–30 years, 31–45 years, 46–60 years), type of habitat (rural, urban) and geographical region (North, Northeast, East, Southcentral, Southwest, Northwest). The prevalence of anogenital HPV infection will be reported by HPV genotype, age groups and geographical region. Second, to examine the concordance between the oral and anogenital samples among men, we will calculate the percentage of participants infected with the same HPV genotype in the oral sample and in at least one anogenital sample. The number and percentage of concordance in infection will be reported for each combination of oral and anogenital samples available. Third, descriptive statistics will be calculated to summarise patients' sociodemographics, HRQOL, health behaviour, sexual behaviour and disease history (table 2). Lastly, to explore risk factors associated with oral and anogenital HPV infection in the target population, multivariate logistic or Poisson regression models will be built. The dependent variables will be oral or anogenital infection and the independent variables will the clinical data, socio-demographics, HRQOL, health behaviour, sexual behaviour and disease history listed in table 2.

DISCUSSION

The risk of cervical HPV infection and cervical cancer in women has attracted great attention from health professionals and the general public. However, the burden of oral HPV infection and related head and neck cancer has been steadily rising in recent decades and requires further investigation.[27] Furthermore, men are often under-represented as an at-risk population for HPV and HPV-related cancers, although men are also affected by HPV and at risk for head and neck and anogenital cancers if infected. In addition, men infected with HPV could transmit the infection to their sexual partners, contributing to the existing burden of HPV and cervical cancer in women. PROGRESS-Plus will generate the much-needed evidence on the HPV burden among men and raise the public awareness of HPV infection among this under-represented at-risk population.

In addition, PROGRESS-Plus will explore risk factors associated with oral HPV and male anogenital HPV infection in mainland China. Identifying important local risk factors, such as sociodemographics and sexual behaviours, is critical to designing and implementing effective public health programmes such as HPV vaccination programmes. This is especially important considering the large geographic (eg, central vs eastern vs Western China, urban vs rural) and cultural differences in these risk factors in China. Results from PROGRESS-Plus will allow healthcare providers, policy makers and other stakeholders to make evidence-based decisions when developing gender-neutral HPV vaccination programmes.

The PROGRESS-Plus study design has many strengths but two main strengths are of the most importance. First, estimating the prevalence of any disease or condition in a country is challenging. To effectively sample the entire population in a country that is geographically large, the sampling approach must follow a multistage selection strategy that stratifies the country by geographical region and type of habitat nested in a region (eg, rural or urban). PROGRESS-Plus uses a two-step sampling design that combines PPS sampling with TCS to cover the entirety of China to select outreach areas and participating healthcare providers within each stratified region, increasing the generalisability of study results. PPS is a method for sampling from a finite population that can improve overall surveillance system accuracy.[28] Since PPS optimises participation without inviting selection bias associated with provider-based studies,[29] recruiting participants using PPS is expected to yield highly accurate information. Second, the processes and approaches established in PROGRESS-Plus will serve as a methodological model or framework for not only future STD-related epidemiological research, but also epidemiological and public health research on other health outcomes in mainland China. These standardised processes and approaches include the design of large-scale epidemiological studies that covers the entire mainland China, community outreach and participant recruitment, and patient data and biological sample collection, management, transportation, analysis, reporting and more. The combination of PPS and TCS with a standardised study protocol ensures the success of PROGRESS-Plus.

PROGRESS-Plus has several potential limitations that are common in epidemiological studies. First, there could be selection bias. People who engage in high-risk sexual behaviour are less likely to participate in medical research because they do not wish that their physician learns about their HPV status or are afraid of a potential positive test result.[30] Second, there could be missing responses in the participant surveys, especially for questions related to sexual behaviour. To minimise missing data, all questionnaires are anonymous and sealed by participants immediately after completion. Data analysts will not know the identity of the participants. Third, our sampling approach is community-based sampling that is not random.

CONCLUSION

PROGRESS-Plus is a multisite, cross-sectional study that aims to examine the prevalence of oral and anogenital HPV infection, the concordance between oral and anogenital HPV infection, and the risk factors associated with these HPV infections in the adult population of mainland China. Results from PROGRESS-Plus will fill important evidence gaps as the study assesses the prevalence of oral HPV infection among both men and women, and the prevalence of anogenital HPV infections among men.

Acknowledgements The authors would like to acknowledge Yi Guo from University of Florida, Anna Giuliano from the Moffit Cancer Center, Laia Alemany from the Catalan Institute of Oncology, Hisham Mehanna from Birmingham University, Haitham Mirghani from Hospital European Georges, and Timothy Waterboer from the German Cancer Center.

Contributors WW conceived and designed the study and drafted the manuscript. QL, YQ and FC assisted in designing the study and writing the manuscript. SK-T, CR, NL, RP, EM and Y-TC assisted in writing the manuscript. All authors read and approved the final manuscript.

Funding The PROGRESS-Plus study is funded by Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, New Jersey, USA.

Map disclaimer The inclusion of any map (including the depiction of any boundaries therein), or of any geographic or locational reference, does not imply the expression of any opinion whatsoever on the part of BMJ concerning the legal status of any country, territory, jurisdiction or area or of its authorities. Any such expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

Competing interests The PROGRESS-Plus study is sponsored by Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Consent obtained directly from patient(s).

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs

Wei Wang http://orcid.org/0000-0002-0855-0002
Fuqiang Cui http://orcid.org/0000-0002-9592-4286

REFERENCES

1 Forman D, de Martel C, Lacey CJ, *et al*. Global burden of human papillomavirus and related diseases. *Vaccine* 2012;30 Suppl 5:F12–23.
2 Kombe Kombe AJ, Li B, Zahid A, *et al*. Epidemiology and burden of human papillomavirus and related diseases, molecular pathogenesis, and vaccine evaluation. *Front Public Health* 2020;8:552028.
3 Castellsagué X. Natural history and epidemiology of HPV infection and cervical cancer. *Gynecol Oncol* 2008;110(3 Suppl 2):S4–7.
4 de Sanjosé S, Brotons M, Pavón MA. The natural history of human papillomavirus infection. *Best Pract Res Clin Obstet Gynaecol* 2018;47:2–13.
5 Bray F, Ferlay J, Soerjomataram I, *et al*. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
6 de Martel C, Plummer M, Vignat J, *et al*. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer* 2017;141:664–70.
7 De Vuyst H, Clifford GM, Nascimento MC, *et al*. Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: a meta-analysis. *Int J Cancer* 2009;124:1626–36.
8 Peder LD de, Silva CM da, Boeira VL, *et al*. Association between human papillomavirus and non-cervical genital cancers in brazil: A systematic review and meta-analysis. *Asian Pac J Cancer Prev* 2018;19:2359–71.
9 Tam S, Fu S, Xu L, *et al*. The epidemiology of oral human papillomavirus infection in healthy populations: a systematic review and meta-analysis. *Oral Oncol* 2018;82:91–9.
10 Mena M, Taberna M, Monfil L, *et al*. Might oral human papillomavirus (HPV) infection in healthy individuals explain differences in HPV-attributable fractions in oropharyngeal cancer? A systematic review and meta-analysis. *J Infect Dis* 2019;219:1574–85.
11 Bruni L, Diaz M, Castellsagué X, *et al*. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *J Infect Dis* 2010;202:1789–99.
12 Sabeena S, Bhat PV, Kamath V, *et al*. Community-based prevalence of genital human papilloma virus (HPV) infection: a systematic review and meta-analysis. *Asian Pac J Cancer Prev* 2017;18:145–54.
13 Hebnes JB, Olesen TB, Duun-Henriksen AK, *et al*. Prevalence of genital human papillomavirus among men in Europe: systematic review and meta-analysis. *J Sex Med* 2014;11:2630–44.
14 Han JJ, Beltran TH, Song JW, *et al*. Prevalence of genital human papillomavirus infection and human papillomavirus vaccination rates among US adult men. *JAMA Oncol* 2017;3:810.
15 Ma X, Wang Q, Ong JJ, *et al*. Prevalence of human papillomavirus by geographical regions, sexual orientation and HIV status in China: a systematic review and meta-analysis. *Sex Transm Infect* 2018;94:434–42.
16 Hang D, Liu F, Liu M, *et al*. Oral human papillomavirus infection and its risk factors among 5,410 healthy adults in china, 2009-2011. *Cancer Epidemiol Biomarkers Prev* 2014;23:2101–10.
17 Zhang C, Liu F, Pan Y, *et al*. Incidence and clearance of oral human papillomavirus infection: a population-based cohort study in rural China. *Oncotarget* 2017;8:59831–44.
18 Wei F, Yin K, Wu X, *et al*. Human papillomavirus prevalence and associated factors in women and men in South China: a population-based study. *Emerg Microbes Infect* 2016;5:e119.
19 Wei F, Li M, Wu X, *et al*. The prevalence and concordance of human papillomavirus infection in different anogenital sites among men and women in liuzhou, China: a population-based study. *Int J Cancer* 2018;142:1244–51.
20 Wei F, Sheng W, Wu X, *et al*. Incidence of anogenital warts in liuzhou, South China: a comparison of data from a prospective study and from the National surveillance system. *Emerg Microbes Infect* 2017;6:e113.
21 He Z, Liu Y, Sun Y, *et al*. Human papillomavirus genital infections among men, China, 2007-2009. *Emerg Infect Dis* 2013;19:992–5.
22 Zhang D-Y, Yin Y-P, Feng T-J, *et al*. Hpv infections among MSM in Shenzhen, China. *PLoS One* 2014;9:e96364.
23 Fan S, Li P, Ouyang L, *et al*. Anal human papillomavirus infection among MSM attending university in China: implications for vaccination. *Vaccines (Basel)* 2020;8:175.
24 Wong MCS, Vlantis AC, Liang M, *et al*. Prevalence and epidemiologic profile of oral infection with alpha, beta, and gamma papillomaviruses in an Asian Chinese population. *J Infect Dis* 2018;218:388–97.
25 Liu F, Hang D, Deng Q, *et al*. Concurrence of oral and genital human papillomavirus infection in healthy men: a population-based cross-sectional study in rural china. *Sci Rep* 2015;5:15637.
26 Hernandez G, Garin O, Dima AL, *et al*. EuroQol (EQ-5D-5L) validity in assessing the quality of life in adults with asthma: cross-sectional study. *J Med Internet Res* 2019;21:e10178.
27 Van Dyne EA, Henley SJ, Saraiya M, *et al*. Trends in human papillomavirus-associated cancers - United States, 1999-2015. *MMWR Morb Mortal Wkly Rep* 2018;67:918–24.
28 Thurmond MC. Conceptual foundations for infectious disease surveillance. *J Vet Diagn Invest* 2003;15:501–14.
29 Treiman DJ, Lu Y, Qi Y. New approaches to demographic data collection. *Chin Sociol Rev* 2012;44:56–92.
30 Kahn JA, Belzer M, Chi X, *et al*. Pre-vaccination prevalence of anogenital and oral human papillomavirus in young HIV-infected men who have sex with men. *Papillomavirus Res* 2019;7:52–61.