BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Kinetics and ecology of Human Papillomavirus (HPV) genital infections in young women: the PAPCLEAR study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025129
Article Type:	Protocol
Date Submitted by the Author:	06-Jul-2018

Complete List of Authors:	Murall, Carmen Lia; Centre National de la Recherche Scientifique (CNRS), MIVEGEC Rahmoun, Massilva; Centre National de la Recherche Scientifique (CNRS), MIVEGEC Selinger, Christian; Centre National de la Recherche Scientifique (CNRS), MIVEGEC Bernat, Claire; Centre National de la Recherche Scientifique (CNRS), MIVEGEC Buisson, Mathilde; Centre Hospitalier Regional Universitaire de Montpellier, Direction Recherche Innovation Christophe, Guillaume; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD D'Auria, Giuseppe; Fundacio per al Foment de la Investigacio Sanitaria i Biomedica, Sequencing and Bioinformatics Service; Centro de Investigacion Biomedica en Red de Epidemiologia y Salud Publica De Taroni, Florence; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Froissart, Rémy; Centre National de la Recherche Scientifique (CNRS), MIVEGEC Hirtz, Christophe; Centre Hospitalier Regional Universitaire de Montpellier, LBPC/PPC, IRMB Jaussent, Audrey; Centre Hospitalier Regional Universitaire de Montpellier, Département de l'Information Médicale Lajoie, Julie; University of Manitoba College of Medicine, Department of Medical microbiology Lorcy, Frédérique; Centre Hospitalier Regional Universitaire de Montpellier, Laboratoire d'anatomie et cytologie pathologiques Picot, Eric; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD PICOT, Marie-Christine; Centre Hospitalier Regional Universitaire de Montpellier, Département d'Information Médicale Ravel, Jacques; University of Maryland School of Medicine, Institute for Genome Sciences Reynes, Jacques; Centre Hospitalier Regional Universitaire de Montpellier, Département des Maladies Infectieuses et Tropicales Rousset, Thérèse; Centre Hospitalier Regional Universitaire de Montpellier, Laboratoire d'anatomie et cytologie pathologiques Seddiki, Aziza; Centre Hospitalier Regional Universitaire de Montpellier, Direction Recherche Innovation Teirlinck, Martine; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD

	Tribout, Vincent; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Tuaillon, Edouard; Centre Hospitalier Regional Universitaire de Montpellier, Pathogenesis and Control of Chronic Infections Waterboer, Tim; Deutsches Krebsforschungszentrum, Infections and Cancer Epidemiology Jacobs, Nathalie; Universite de Liege Faculte des Sciences, GIGA-Research, Cellular and molecular immunology Bravo, Ignacio; Centre National de la Recherche Scientifique (CNRS), MIVEGEC Segondy, Michel; Centre Hospitalier Regional Universitaire de Montpellier, Pathogenesis and Control of Chronic Infections Boulle, Nathalie; Centre Hospitalier Regional Universitaire de Montpellier, Pathogenesis and Control of Chronic Infections Alizon, Samuel; Centre National de la Recherche Scientifique (CNRS), MIVEGEC
Keywords:	VIROLOGY, IMMUNOLOGY, Epidemiology < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts

STUDY PROTOCOL

Kinetics and ecology of Human Papillomavirus (HPV) genital infections in young women: the PAPCLEAR study

Carmen Lía Murall[1], Massilva Rahmoun[1], Christian Selinger[1], Claire Bernat[1], Mathilde Buisson[2], Guillaume Christophe[3] , Giuseppe D'Auria[4,13], Florence De Taroni[3], Rémy Froissart[1], Christophe Hirtz[5], Audrey Jaussent[6], Julie Lajoie[7] , Frédérique Lorcy[14], Eric Picot[3], Marie-Christine Picot[6], Jacques Ravel[9], Jacques Reynes[10], Thérèse Rousset[14] , Azziza Seddiki[2], Martine Teirlinck[3], Vincent Tribout[3], Édouard Tuaillon[8], Tim Waterboer[11], Nathalie Jacobs[12] , Ignacio G Bravo[1], Michel Segondy[8], Nathalie Boulle[8], Samuel Alizon[1*]

Word count: 4860 words, excluding title page, abstract, references, figures and tables.

* Corresponding author: samuel.alizon@cnrs.fr , phone: +33 6 10 65 49 02

1 Laboratoire MIVEGEC (UMR 5290 CNRS, IRD, UM), 911, avenue Agropolis, BP 64501, 34394 Montpellier, France

2 Direction Recherche Innovation, Centre Hospitalier Universitaire de Montpellier, Montpellier, France

3 CeGIDD, Centre Hospitalier Universitaire de Montpellier, Montpellier, France

4 Sequencing and Bioinformatics Service, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública), 46020 Valencia, Spain

5 University of Montpellier, LBPC/PPC- IRMB, CHU de Montpellier, 80 rue Augustin Fliche, Montpellier, France

6 Département de l'Information Médicale, Centre Hospitalier Universitaire de Montpellier, Montpellier, France

7 Department of Medical microbiology, University of Manitoba, 745 Bannatyne, Winnipeg, Canada

8 Pathogenesis and Control of Chronic Infections, INSERM, CHU, University of Montpellier, Montpellier, France

9 Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, USA

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

10 Département des Maladies Infectieuses et Tropicales, Centre Hospitalier Universitaire de Montpellier, Montpellier, France

11 German Cancer Research Center (DKFZ), Infections and Cancer Epidemiology, Im Neuenheimer Feld 280, Heidelberg, Germany

12 GIGA-Research, Cellular and molecular immunology, University of Liège, 3 Avenue de l'Hôpital, 4000 Liège, Belgium

13 CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain

14 Laboratoire d'anatomie et cytologie pathologiques, Hôpital Gui de Chauliac, Centre Hospitalier Universitaire de Montpellier, 80 avenue Augustin Fliche, 34295 Montpellier, France

For peer review only

ABSTRACT

Introduction

Human papillomaviruses (HPVs) are responsible for one third of infection-induced cancers. However, most HPV studies focus on chronic infections and cancers, and we know little about the early stages of the viral infection. In particular, the roles of the immune system, the microbiota, the virus genetics and of the host genetics on infection clearance or persistence remains poorly understood.

Methods and Analysis

We follow 150 women aged 18-25 longitudinally to monitor immune response features (cytokines and immune cells in the genital tract, circulating anti-HPV antibodies), HPV virus load and vaginal microbiota composition. This is complemented by virus and human genetics and behavioural data. To increase the statistical power for the epidemiological framework, an additional 150 women are screened cross-sectionally.

Ethics and Dissemination

This study will provide us with one of the most detailed follow-up studies of acute HPV infections and their interactions with the host and the vaginal microbiota. It will also allow us to investigate related issues regarding HPV intra-host evolution and diversity, vaginal microbiota dynamics and sexually transmitted infections. The trial has been registered to ClinicalTrials.gov on 27 Oct 2016 with ID number NCT02946346.

For peer review only

ARTICLE SUMMARY

STRENGTHS AND LIMITATIONS OF THIS STUDY

- We set up a longitudinal study to investigate the natural history of HPV genital infections in N=150 young women.
- The follow-up is dense (visit every 2 month for infected women) and long (up to 24 months).
- At each visit, the goal is to estimate virus load, cytokine densities, immune cell counts in the cervical area and vaginal microbiota composition.
- Clinical data will be combined to population dynamics models to perform parameter estimation and model comparison.
- The longitudinal study is combined with a cross-sectional study of N=150 women to allow for epidemiological analyses.

For peer review only

Keywords: HPV; acute infection; persistence; virus load; immunity; microbiota; population dynamics

INTRODUCTION

Epidemiology of HPV genital infections in young adults and public health implications

Infections by Human Papillomaviruses (HPVs) are probably the most common sexually transmitted infections (STI) globally. It is often estimated that ≈ 20% of men and women in European countries are currently infected by HPVs and that, worldwide, 75% of the individuals will be infected at some point in their life by HPV [1]. In France, a recent study performed in the Paris area estimated prevalences of genital HPV of ≈ 16% in women, and ≈ 25% in women below 25 years of age [2]. In the area of Montpellier, prevalence of oncogenic HPVs (often referrer to as 'high-risk', HR) in pregnant women aged 16 to 42 was close to 20% [3].

Fortunately, the vast majority of HPVs infections are asymptomatic and benign. Even for HPV16, probably the most oncogenic biologic agent to humans, only a minority (less than 10%) of infections becomes persistent [4], and again a minority of these (12%) will progress into cancer if untreated [1, 5]. Indeed, it is estimated that approximately 70 to 100% of HPV infections are cleared within 12 to 24 months, even for the most oncogenic HPVs (such as HPV16 and 18) that are responsible for the majority of HPV-induced cancers [1, 4, 6, 7]. Yet, we currently know little about the biology of these very prevalent non-persisting infections [8].

Our lack of knowledge partly comes from the fact that hitherto studies interested in persistent infections follow participants every six months for several years (e.g. a median 50.4 months in [9]). This is indeed sufficient to assess the time to clearance (or to persistence) but it is clearly not precise enough to understand the kinetics of infections that last on average between six to 24 months. After 24 months of infection, an infection is often considered as being persistent [10].

We know some factors that correlate with persistence (e.g. immunosuppression, smoking, and co-infection with other STIs [11]) but we do not know how these play out in the within-host dynamics of infections. Also, there are hypothesized changes in viral-immunity interactions that appear related to persistence and disease progression [12–15] but, again, we do not know the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

underlying interactions between the viruses, the host target cells and the immune response in acute infections [8]. Finally, it has been argued that the vaginal microbiota may differs between HPV-infected and HPV-uninfected women [16] and that specific microbiota composition may interact with HPV detection [17]. However, it is difficult to disentangle the cause and the consequence. For instance, does the microbiota composition change after the establishment of an HPV infection, or do certain microbiota compositions increase susceptibility to HPV infection?.

A better understanding of the kinetics of HPVs infections and of the determinants of clearance and persistence of viral infection is particularly important in the context of vaccination [18–21]. Indeed, the long-term efficacy of the vaccine at the population level will largely depend on the virus within-host dynamics. Furthermore, a better understanding of acute HPVs infections can shed a new light on issues related to latency, fertility, or immunotherapies [8].

Prevention strategies and treatment

Treatment

Since the majority of HPVs infections are benign in young adults and clear within six to 24 months, the current standard of care is to avoid over-treatment [22]. Clinical interventions (colposcopies, biopsies and treatment) are performed less in young women (< 25) and only for high-grade (pre-cancerous) lesions (cervical intraepithelial neoplasia grade 2, CIN-2, or more). Low-grade lesions (CIN-1) are not systematically treated but rather followed (approximately every twelve months) to detect any progression to high-grade lesions.

Genital warts caused by non-oncogenic HPVs (often referred to as ‘low-risk’ (LR) HPVs) can be removed by surgery or treated with bi- and trichloroacetic acid, cryotherapy or other treatments [23].

HPV vaccination

There are currently three licensed vaccines: a bivalent vaccine (Cervarix) targeting HPV16 and HPV18 (the most oncogenic and the most prevalent HPVs), a quadrivalent vaccine

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(Gardasil®) that additionally targets HPV6 and HPV11 (non-oncogenic, but highly prevalent and associated to benign proliferative lesions) and, since 2014, a nonavalent vaccine (Gardasil 9®) that targets five more oncogenic types, HPV31, HPV33, HPV45, HPV52, and HPV58. These vaccines succeed in eliciting a protective immune response against new infections by the targeted viruses, and are used throughout the world with wide variation in coverage (for reviews, see e.g. [24, 25]).

In France, vaccination started in 2006 but with limited coverage: it was 28.5% in 2008 [26] and has been decreasing since then [27]. The vaccine is recommended for girls from 11 to 14 (current vaccination scheme is two doses with a six months interval), and with a catch-up for girls aged 15-19 (three doses). It is reimbursed by the social security but not mandatory. It is also recommended for men who have sex with men (MSM) and for immunocompromised people [27]. Vaccination is now the primary prevention strategy against cervical cancers.

Screening

In France, the secondary prevention strategy against cervical cancer is routine individual cytology-based screening for pre-cancerous and cancerous cervical lesions in women between 25 and 65 years. Cytologies can also be performed in younger women if they report risk factors for cervical cancer (multiple partners, chronic STIs or HIV status [27]). Detection of oncogenic HPVs is proposed for triage in case of abnormal cytology (i.e. ASCUS, or Atypical Squamous Cells of Undetermined Significance).

Primary objectives

The first objective is to decipher the kinetics and ecology of HPVs cervical infections, i.e. the population dynamics of the virus, target epithelial cells and immune effectors in healthy young women.

The second objective is to determine the prevalence of genital HPVs in young women in the region of Montpellier, in relationship with lifestyle, vaginal microbiota and human genetics.

For peer review only

Secondary objectives

A secondary objective is to characterize the acquisition and clearance dynamics of cervical HPVs infections as a function of viral diversity, host immunity, vaginal microbiota and human genetics.

Finally, another secondary objective is to investigate the genetic diversity of HPVs during cervical infections.

METHODS AND ANALYSIS

Participants

The study population is young women aged 18 to 25 at risk of HPV infection. The estimated prevalence in this age class is approximately 25% [2], and it decreases to 15% in women older than 25.

The composition of the population visiting the Montpellier STI detection centre (CeGIDD) has already been documented in an earlier study [28]. In total, the centre is visited by approximately 3,000 women per year, who tend to be less than 25 years old (80%). Approximately 40% of the attendants report three or more partners over the last twelve months and approximately 50% report using adopting adequate behaviour for prevention against HIV. Overall, in terms of STI exposure, the centre is equally visited by people with high-risk and low-risk behaviours.

Inclusion criteria

Participants are women aged 18 to 25. They must be sexually active with at least one new partner over the last 12 months. This criteria is fixed to maximise the incidence of new HPVs infections. As in any clinical study, participants must be able to and willing to give written informed consent: they must sign and informed consent, understand the requirements for the study and be affiliated to a social security scheme.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Women cannot be included in the study if they have a history of cervical pathology (genital warts or cervical lesions), if they are pregnant or intending to become pregnant in the coming year, infected by HIV, undergoing (or planning to undergo) heavy treatment (biotherapy, chemotherapy, immunosuppression), planning on moving outside the Montpellier metropole within the next 18 months, in a dependency or employment with the sponsor or the investigator, if they participated in a clinical trial involving administration of drugs within the last four weeks or if they belong to a vulnerable group (guardianship).

Design/setting

Our study has a longitudinal component aimed at deciphering within-host kinetics and a cross-sectional component, aimed at understanding the epidemiology of HPVs infections in young adults in the area of Montpellier, France. The general structure of the study is shown in Figure 1.

If participants fit the inclusion criteria, they go through an inclusion visit (V_1) with a gynaecologist, at the CeGIDD. During this visit, they fill out health and lifestyle questionnaires and undergo a medical consultation during which a number of samples are collected. Participants are given self-sample swabs for collecting microbiome samples until the next visit and are instructed on how to fill out weekly questionnaires through an online form (these are performed throughout the study).

An appointment is scheduled four weeks later for the results visit (V_2), where the cervical cytology results are communicated. We collect some samples and provide more self-sample swabs for home collection.

The next return visits (V_i, where $i > 2$) are as follows:

- Participants infected by an *Alphapapillomavirus* at V_1 join the HPV positive (HPV+) arm of the study with return visits scheduled every 2 months.
- Participants who were not infected by an *Alphapapillomavirus* at V_1 join the HPV negative (HPV-) arm with return visits scheduled every 4 months.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

- HPV- participants who become infected by HPV move to the HPV+ arm.

Participants in the HPV- arm will be followed until month 26 of the study.

Participants in the HPV+ arm will be followed until they clear the infection or until they have been infected for 24 months (after which we consider that the infection is persistent). Clearance is defined as being negative two visits in a row for the first HPV type detected in the follow-up.

In between these visits to the CeGIDD, participants are asked to perform regular (every week for HPV+ and every second week for HPV-) self-samples using vaginal swabs, along with a measure of vaginal pH and filling a short questionnaire.

Patients and public involvement

Participants are all healthy and are therefore referred to as such rather than patients. They will be mainly recruited amongst the people visiting the CeGIDD via poster, leaflet and direct contact with the Clinical Research Technician (TRC) or the clinicians. To maximise inclusion, recruitment will also target students from the various Universities in Montpellier. Finally, a social media page will be set up.

Participants did not play a role in the design of this study.

Results of the study will be disseminated to study participants via email.

Visits

Inclusion visit (*V*1)

This visit takes place at the CeGIDD and is scheduled by the Clinical Research Technician (TEC) via phone or email.

Women meet a physician investigator, who explains the study goal and requirements. The

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

physician also checks that the inclusion criteria are met. If so, after a general discussion, the informed consent forms are signed.

The physician first performs a general exam, before performing a gynaecological exam during which the following samples will be taken:

- vaginal pH cotton swab (EcoCare™)
- vaginal swab (Copan ESwab™) in 1mL Amies liquid for DNA extraction (microbiota analysis)
- vaginal swab (Copan ESwab™) in 1mL of RNA preservation medium
- ophthalmic sponge (Weck-cel□) to collect cervical secretions (cytokine analysis)
- cervical smear in 20mL of Thinprep□ Preservcyt□ fixation liquid (HPV and HSV assays, and cytology evaluation)

Following the gynaecological consult, the participant meets with a nurse to measure body temperature, blood pressure and draw 20mL of blood: a 5mL tube for SNPs sequencing, a 10mL tube for immunophenotyping and a 5mL tube for HPV antibody titration. For the longitudinal study, the nurse provides the participant with 3 self-sampling kits, 3 pH strips, a freezer box to bring back to the next visit, and instructions on how to perform the home sampling.

If the participant has not been tested for a STI in the last 3 months, the nurse will draw an additional blood tube of 5mL to test for STIs (HIV, HCV, HBV) and ask for self-samples for chlamydiae and gonorrhea detection. Syphilis testing was also prescribed based on the STI clinic's standards.

Finally, the participant meets with the TEC to fill in questionnaires #1 (inclusion visit) and #3 (home). The TEC answers any remaining questions, explains how to fill the home questionnaires (#3) and sets an appointment for the results visit.

For peer review only

Results visit (*V*2)

During this visit, the participants are informed if the analysis of the liquid cytology indicated a cervical lesion (ASCUS, LSIL or HSIL). Participants with a high grade lesion (HSIL) exit the study and are referred to the gynaecology service of the CHU of Montpellier.

During this visit, the gynaecologist collects additional samples: 2 vaginal swabs for DNA and RNA analysis, and a cervical smear in 10mL of PBS to confirm HPV status and perform flow cytometry analyses (FACS).

The participant fills in questionnaires #2 (for return visits) and #3 (home). An appointment for the next visit is set and additional home self-samples are given.

Return visits (*Vi*)

These visits only occur in the longitudinal study.

HPV- arm Participants uninfected by HPV visit the **clinic** every 4 months until month 26. During these visits, the same samples as in the inclusion visit (V_1**)** are collected by the gynaecologist.

The nurse only draws blood if a screening test for other STIs than HPV is required. The participant then fills in questionnaires #2 and #3 and an appointment is set for the next visit in 16 weeks.

If an HPV infection is detected based on the samples collected during this visit, the TEC will contact the participant to move the appointment forward.

HPV+ arm Participants infected by HPV visit the clinic every 2 months until clearance or chronic infection. During these visits, the same samples as in the inclusion visit (V_0) are collected during the gynaecological exam.

Then the nurse draws 5mL of blood for HPV antibody titrating. If this is the first HPV+ visit following an HPV- visit, the nurse draws an additional 10mL of blood for immunophenotyping. If a test for additional STIs is needed, the nurse draws 5mL more of blood and asks for a self-

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

sample for STI detection.

Importantly, if the participant has been infected by a HR-HPV for more than 12 months and a cytology has not been performed within the last 12 months, the cervical smear is put in Thinprep® fixation medium, instead of PBS, for cytological analysis (cervical lesion screening).

Finally, the participant will fill in questionnaires #2 and #3, receive self-samples for home collection and an appointment is set for the next visit in 8 weeks.

Endpoints

The primary endpoint for the study is the kinetics of the HPV virus load, associated to local cytokines profile, local immune cells and cervical smear cytology.

Secondary endpoints are the interaction between the course of the infection (e.g. duration), the HPV type(s), the bacterial, fungal and viral communities in vaginal microbiota, human genetics (SNPs) and basal immunological status.

Technical procedures

DNA extraction

DNA extraction from cervical smears will be performed using Nuclisens EasyMAg from Biomerieux or an equivalent protocol. For the microbiota analyses, special kits involving physical (via beads) and/or enzymatic breaking of the cellular barrier will be favoured following standard protocols to study the vaginal microbiome [29], e.g. the MagAttract® PowerMicrobiome® DNA/RNA kit from Qiagen. Detection will be based on 16S RNA loci for the bacteria and ITS loci for fungi. We anticipate that the bacteria should belong to the OTU described in the five community state types [30, 31].

HPV detection, typing and quantification

The participants' infection status (HPV+ or HPV-) will be assessed using the DEIA test, which is based on a PCR of the short SPF10 amplicon [32] and detects all *Alphapapillomaviruses* with great sensitivity.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

If the DEIA test is positive, the present viruses will be typed using the LiPA25 technique, which is based on the same SPF10-PCR, and argued to have a lower detection threshold than other hybridisation-based typing methods [33].

The reason for basing the detection on the DEIA rather than the LiPA25 is that some *Alphapapillomavirus* may not be detected by DEIA but not genotyped by LiPA and also that the DEIA is more sensitive than the LiPA. If the DEIA is positive and the LiPA25, we will sequence part of the HPV genome using a PCR targeting another region of the genome that the SPF10 (e.g. PGMY09/11 [34]) to determine which type it belongs to

The quantification of HPV DNA genome copy number in the samples will be performed using the protocol set up by [35].

Cytokine titration

Cervical sponges will be centrifuged after the addition of 200μL of PBS. Cervical secretions will be analyzed for a set of 5 to 6 cytokines levels using the Meso Scale Discovery (MSD) Multiplex ELISA platform, which allows a low detection threshold and a slowly saturating dose-response curve. A large spectrum of cyctokines will be explored first to choose the most relevant ones (see also [36, 37]).

Flow cytometry

Analysing immune cells via flow cytometry is extremely challenging on cells as fragile as the ones from cervical smears. However, several studies suggest that this is feasible [36–38]. Here, we will follow the protocol described in [39].

Stainings will be performed using a Duraclone custom mix targeting CD45, CD3, CD4, CD8, CD16, CD56, CD69, CD161 and TCR$\gamma\delta$. The last marker, Live&Dead will test for cellular viability. Samples will be acquired using a Navios flow cytometer (Beckman Coulter, three-laser configuration).

For peer review only

Sequencing

Sequencing will be performed for microbiota profiling. It will involve PCR amplification of 16S RNA (for bacteria) and ITS (for fungi). The virome will also be explored using shotgun sequencing. Human genetics will be explored using chip sequencing for SNPs.

Statistical analyses

Times series analyses

The core results of the study will come from the longitudinal follow-up of infected women, which will generate time series, i.e. a set of values collected from the same individual over time (Figure 2). There will be several time series per individual (virus load, number of immune cells, cytokine and antibody levels). These time series will be used to fit mathematical population dynamics models that describe the interaction between viruses, host target cells (here, in the case of HPV, keratinocytes) and the immune response. These models are commonly developed for viral infections [40–42] and are being developed for HPV [43].

We will use non-linear mixed effect models [44] to jointly analyse time series from all participants. More precisely, we will rely on *R* packages such as nlme [45] or lme4 [45]. Note that in addition to estimating model parameters (e.g. life-expectancy of infected cells or virion production rate of infected cells), this approach can also allow us to compare biological models using statistical tools based on model likelihood such as Akaike Information Criterion. For an example of such analysis in the case of HIV, see [41].

Microbiota dynamics

The composition of the vaginal microbiota has already been described and shown to exhibit much less diversity than the gut microbiota for instance [30]. The dynamics of this microbiota has also been studied and shown to closely follow menstrual cycles [31].

We will use the time series of OTU abundances (measured via 16S RNA sequencing and qPCR) to infer interaction parameters by assuming an underlying Lotka-Volterra competition model [46]. This work will include time series analysis techniques (e.g. auto-correlation or local

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

similarity analysis) and statistical inference methods in order to infer community structure and interactions from the NGS datasets [47]. Finally, statistical methods from ecology will also be used to study community diversity (e.g diversity indices) and community assembly, such as cluster and ordination analyses [48].

Genome Wide Association Studies

In our analysis, we will use human single nucleotide polymorphisms (SNPs) inferred by chip sequencing to look for genetic determinants of key traits (e.g. microbiota composition, HPV infection duration). This is classically done by performing a Genome Wide Association Study (GWAS), which is a complex regression method designed for situations where there are many explanatory variables (here millions of SNPs) for a single trait of interest. GWAS will be performed using classical methods [49]. Earlier GWAS studies have been applied to HPV infections for instance to test for determinants to the ability to seroconvert following infection [50] and cervical cancer (see [51] for a review).

Additional analyses

For all collected variables, descriptive statistics will be calculated according to the level of measurement. For metric variables this contains mean and standard deviation as well as median and range of the data. In case of categorical variables group proportions and contingency tables are prepared.

Univariate inferential statistics will follow the descriptive analysis. Generally, parametric testing procedures are preferred to non-parametric tests, as the former have higher power. That is why, for metric variables, a check whether the data can be assumed normally distributed will be first conducted. For normally distributed variables, ANOVA statistics will be done to detect differences between the groups. In case of significance, post-hoc analysis (Tukey test) are planned to reveal pairwise differences. If the data are not normally distributed or ordinally scaled, non-parametric analysis will be used. This contains the Kruskal-Wallis test and the Wilcoxon test as a post-hoc test with an appropriate correction of the significance level. Since the cell counts are expected to be small, Fisher's exact test will be performed for contingency

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

tables instead of the asymptotic chi-square test for categorical variables.

Sample size calculation

The study will enrol a total of *N* = 300 women, with *N* = 150 in a longitudinal study and *N* = 150 in a cross-sectional study. The goal of the longitudinal study is to follow approximately 75 and at least 40 women longitudinally, preferentially before they are infected (see above).

For this, 150 participants will be enrolled longitudinally in the study. Enrolment will stop if 75 infected women are being followed. Dropouts of infected women during the enrolment period (i.e. until month 22) will be replaced For the following calculations, we assumed a high percentage of lost during follow-up (30%).

With 150 enrolments and considering that the prevalence of HPV infection in the young women is ≈ 60% (based on our preliminary data) and 30% of lost to follow-up, we expect to detect (and successfully follow) 63 infections at inclusion [CI95: 51−75], using a 95% confidence interval assuming a binomial distribution.

Among the uninfected women at the first visit and considering the yearly incidence being close to 30% [52] and 30% of lost to follow-up, we expect 12 [CI95: 6−20] to be infected during the first year of follow-up.

In the end, with 150 enrolments and assuming a high percentage of lost to follow-up (30%), we expect to successfully follow 75 [CI95: 56−95] women infected at different stages of HPV infection: beginning, during and end.

Note that this will be made possible by the probability of transmission of HPV, which is extremely high without condom use (≈ 90%) and still high with condom use (≈ 40%) [53]. Moreover, only ≈ 50% of the target population at the CeGIDD reported adopting safe-sex prevention measures [28].

Finally, regarding potential interference with the HPV vaccines, we do not anticipate any significant problem for two reasons. First, as mentioned above, the vaccine coverage is low in

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

France. Second, and more importantly, the vaccines only target few HPV types, thus leaving open the possibility of infection by dozens of types. Furthermore, studying the kinetics of a non-vaccine HPV type in a vaccinated woman will be extremely informative.

To run cross-sectional analyses (especially on the microbiota and human genetics), we will enrol *N* = 150 women who will only perform the inclusion and the results visit. This sample size is based on that of some earlier GWAS studies [51].

Trial governance

Sponsor

This study is sponsored by the Centre Hospitalier Universitaire (CHU) of Montpellier. The CHU is involved in the implementation of the trial, legal/ethical submissions and implementing the database (eCRF), which is hosted by Ennov-Clinical (ClinSight). The CHU is not involved in the analysis or interpretation of the data. The CHU of Montpellier performs regular quality control assessments. A clinical research assistant will visit the CeGIDD every 4 months to ensure that implementation is in accordance with the protocol. The CHU has taken out insurance from the Société hospitalière d'assurances mutuelles, 18, rue Edouard Rochet-6 9372 Lyon cedex 08 (contract number 138983) through the full research period, covering its own civil liability and that of any agent (doctor or research staff), in accordance with article L.1121-10 of the French Public Health Code.

Scientific committee

The scientific committee comprises the study investigators, clinicians, scientific experts and representatives of the sponsor. It will meet yearly. It will be responsible for following research progress, monitoring compliance with good clinical practice and patient safety. It will also be able to decide relevant modification of the protocol. Request from third parties to access data collected during the study will be evaluated by the committee.

For peer review only

Monitoring

Monitoring will be performed during the whole study at CeGIDD according to the sponsor specific SOP. Routine monitoring visits will be made by the monitors designated by the sponsor to check compliance with the protocol, the completeness, accuracy and consistency of the data, and adherence to GCP. The principal investigator must ensure that eCRFs are completed in a timely manner and must allow periodical access to eCRFs, patient records, drug logs and all other study-related documents and materials. The frequency of monitoring visits will be determined by factors such as study design and the site enrolment requirements but visits will normally occur at least once every 4 months.

Trial registration

The trial has been registered to ClinicalTrials.gov on 27 Oct 2016 with ID number NCT02946346.

DISCUSSION

Expected results

HPVs acute infections have been put into the spotlight because vaccination is much more efficient when it occurs before primo-infection. However, we currently know very little about the early stages of HPV infections. This clinical study will give us an unprecedented level of detail on the kinetics of HPVs infections in young women. Variations in virus load have been studied but in the context of cervical cancer in older women [54]. In addition to variations in virus load, we will also have access to the description and the dynamics of parts of the immune response (local immune cells and cytokines, circulating anti-HPV antibodies) and of the vaginal microbiota. Beyond these kinetics, we will also have access to other informations regarding the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

infection such as clearance or not in 24 months, presence of more than one HPV type or coinfection by other STIs.

To analyse these data, we will have access to numerous cofactors. One of the most important will be human genetics, with the sequencing of millions of SNPs. Others will be related to the sexual behaviour (number of partners, contraception methods, sexual practices) and general life. We therefore expect general insights regarding sexual health in young women.

Practical and operational issues

Practically, one of the main challenges resides in the analysis of cervical smears by flow cytometry. Indeed, the tissues are known to be fragile, adhesive and auto-fluorescent. Even though standard protocols now exist [39], they require the process of fresh samples in less than 2 hours.

Another potential issue has to do with contaminations, which are frequent in the HPV field due to the robustness of the virions and the sensitivity of the tests. To certify our ability to control for these, we have entered the 2017 GLOBAL HPV DNA Proficiency Panel from the WHO HPV LabNet [55].

Regarding the enrolment of the participants, given the number of visitors of the centre who fit the inclusion criteria (more than 3,000 per year) and given earlier high participation rates in the same population ([28] enrolled 1381 participants in 5 months for their study) we do not expect to meet any problems to enrol 150 women in 22 months for the longitudinal study and 150 for the cross-sectional study.

As in any longitudinal study, ensuring participant commitment will be challenging. To achieve this goal, we have set up a compensation of 40 EUR per visit and an additional 10 EUR times the total number of visits in case of a complete follow-up. Furthermore, participants who have answered a sufficient number of questionnaires and brought back a sufficient number of self samples will get a 100 EUR bonus at the end. Overall, a participant performing 12 return visits would gain a total compensation of 650 EUR.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Concerning the follow-up, the high incidence rate of HPV can also lead to transient carriage, i.e. women who are positive for a type only at a single visit. This has been observed for instance in longitudinal studies with a tight follow-up interval [17]. To control for this, we will run the HPV detection test on the cells from the cervical smear after washing with RPMI.

For peer review only

Table 1: **Summary of the visits schedules and sample taken**. The cross-sectional study only includes the first two columns (V_1 and V_2). The · indicate samples taken at visits. □ participants infected by a HR-HPV for 12 month will have one PBS smear replaced by a Thinprep□ smear to perform a cytology and check for lesions. □ this sample is only taken at the first HPV+ visit of a formerly HPV- participant. □ STI detection will be performed at inclusion unless the participant has been tested within the last 3 months and during the study every 6 months if a new partner has been reported or upon request.

For peer review only

	Inclusion (V_1)	Results (V_2)	Return (V_i, with $i > 2$)	
Participants	all	all	HPV+	HPV-
Time	day 0	+ 4 weeks	+ 8 weeks	+ 16 weeks
Eligibility	·			
Consent	·			
Gynecological consult	·	·	·	·
Vaginal pH coton swab	·	·	·	·
2 vaginal swab samples (Copan ESwabTM)	·	·	·	·
1 ophtalmological sponge sample	·		·	·
1 cervical smear in Thinprep□ (cytology)	·		□	
1 cervical smear in PBS		·	□	·
Blood sampling (HPV antibodies)	·		·	
Blood sampling (sequencing)	·			
Blood sampling (immunophenotyping)	·		□	
Other STI detection	□	□	□	□
Questionnaire #1 (inclusion)	·			
Questionnaire #2 (visit)		·	·	·
Questionnaire #3 (home)	·	·	·	·
Returning self-sampling samples		·	·	·
Serious Adverse Event collection		·	·	·

Abbreviations

ANOVA: Analysis of variance,

ASC-US: Atypical squamous cells of undetermined significance,

CD: Cluster of differentiation,

CI95: 95% Confidence interval,

CeGIDD: :Centre Gratuit d'Information de Dépistage et de Diagnostic,

CHU: Centre Hospitalier Universitaire,

CIN: Cervical intraepithelial Neoplasia,

ELISA: enzyme-linked immunosorbent assay,

GWAS: Genome Wide Association Study,

HIV: Human Immunodeficiency Virus,

HPV: Human Papillomavirus,

HR: high-risk,

ITS: Internal Transcribed Spacer,

HSIL: High grade Squamous Intraepithelial Lesion,

LR: low-risk,

LSIL: Low grade Squamous Intraepithelial Lesion,

OTU: Operational Taxonomic Unit,

PBMC: Peripheric Blood Mononuclear Cell,

PBS: Phosphate Buffered Saline,

RPMI: Roswell Park Memorial Institute medium,

SNP: Single Nucleotide Polymorphism,

TCR: T-cell receptor,

WHO: World Health Organisation.

TRIAL STATUS

The study began on Oct 1, 2016 and the first inclusion was on Nov 3, 2016. On Jun 23, 2018,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

89 participants have been included in the longitudinal study. Inclusions in the longitudinal study will continue until Dec 2018 and the study is expected to last until Nov 2020.

CONFLICTS OF INTERESTS

The authors have read and understood BMJ policy on declaration of interests and declare that they have no competing interests.

For peer review only

FUNDING

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 648963 to SA). SA acknowledges further support from the Centre National de la Recherche Scientifique (CNRS) and the Institute de Recherche pour le Développement (IRD).

DATA STATEMENT

All personal and identifying information collected from participants are kept in a secure place at the CeGIDD during the duration of the trial and will be destroyed at the end of the study. The final raw dataset will be accessible only by the sponsor (CHU) and the chief scientist's (SA) team. Anonymous data will be available to external parties upon approval of both the sponsor and the scientific committee. All publications will be made green or gold open access and the corresponding data will be provided as supplementary material or via a public repository (e.g. Dryad), provided that there is no conflict with ethical guidelines.

AUTHOR CONTRIBUTIONS

SA, CLM and MR were the major contributors in the conception of the protocol. All authors were involved in the conception of the protocol or in the implementation of the trial. SA wrote the

initial version of the manuscript and NB, CB, JL, MR, CLM and CS further edited it. All authors read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The PAPCLEAR trial obtained favourable opinions from the Comité de Protection des Personnes (CPP) Sud Méditerranée I on May 11, 2016 (CPP number 16 42, reference number ID RCB 2016-A00712-49); from the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (CCTIRS) on July 12, 2016 (reference number 16.504); and from the Commission Nationale Informatique et Libertés (CNIL) on Dec 16, 2016 (reference number MMS/ABD/AR1612278, decision number DR-2016-488). This trial was authorised by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM) on July 20, 2016 (reference 20160072000007).

The protocol has been modified since its initial version and the modification was submitted to the CPP on Jan 29, 2018. In case the protocol needs to be further modified, the investigator-coordinator will submit a request to the CPP and send an information note to all the investigators.

All participants in the study will sign an informed consent form prior to participation.

For peer review only

ACKNOWLEDGEMENTS

We thank all the study participants and the CeGIDD staff. The EVOLPROOF advisory committee is composed of Giuseppe d'Auria, Michael Blum, Ignacio G Bravo, Marc Choisy, Anne Cori, Elisabeth Delarocque-Astagneau, Christophe Depuydt, Rémy Froissart, Jérémie Guedj, Robert D Holt, Nathalie Jacobs, Julie Lajoie, Dorothée Missé, Lulla Opatowski, Vincent Pedergnana, José-Miguel Ponciano, Jean-Luc Prétet, Jacques Ravel, François Renaud, Jacques Reynes, Silvia de San José, Catherine Tamalet, Javier Tamames, Muriel Thomas, Édouard Tuaillon, Tim Waterboer.

For peer review only

REFERENCES

1. Tota, J.E., Chevarie-Davis, M., Richardson, L.A., Devries, M., Franco, E.L.: Epidemiology and burden of HPV infection and related diseases: implications for prevention strategies. Prev Med **53 Suppl 1**, 12–21 (2011). doi:10.1016/j.ypmed.2011.08.017

2. Monsonego, J., Zerat, L., Syrjänen, K., Zerat, J.C., Smith, J.S., Halfon, P.: Prevalence of genotype-specific HPV infection among women in France: implications for screening and vaccination. Gynecol Obstet Fertil **41**(5), 305–313 (2013). doi:10.1016/j.gyobfe.2013.03.003

3. Brun-Micaleff, E., Coffy, A., Rey, V., Didelot, M.-N., Combecal, J., Doutre, S., Daurès, J.-P., Segondy, M., Boulle, N.: Cervical cancer screening by cytology and human papillomavirus testing during pregnancy in french women with poor adhesion to regular cervical screening. J Med Virol **86**(3), 536–45 (2014). doi:10.1002/jmv.23764

4. Insinga, R.P., Dasbach, E.J., Elbasha, E.H., Liaw, K.-L., Barr, E.: Incidence and duration of cervical human papillomavirus 6, 11, 16, and 18 infections in young women: an evaluation from multiple analytic perspectives. Cancer Epidemiol Biomarkers Prev **16**(4), 709–15 (2007). doi:10.1158/1055-9965.EPI-06-0846

5. Woodman, C.B.J., Collins, S.I., Young, L.S.: The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer **7**(1), 11–22 (2007). doi:10.1038/nrc2050

6. Rodríguez, A.C., Schiffman, M., Herrero, R., Wacholder, S., Hildesheim, A., Castle, P.E., Solomon, D., Burk, R., Proyecto Epidemiológico Guanacaste Group: Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections. J Natl Cancer Inst **100**(7), 513–7 (2008). doi:10.1093/jnci/djn044

7. Trottier, H., Mahmud, S., Prado, J.C.M., Sobrinho, J.S., Costa, M.C., Rohan, T.E., Villa, L.L., Franco, E.L.: Type-Specific Duration of Human Papillomavirus Infection: Implications for Human Papillomavirus Screening and Vaccination. J Infect Dis **197**(10), 1436–1447 (2008). doi:10.1086/587698

For peer review only

8. Alizon, S., Murall, C.L., Bravo, I.G.: Why Human Papillomavirus Acute Infections Matter. Viruses **9**(10), 293 (2017). doi:10.3390/v9100293

9. Herrero, R., Wacholder, S., Rodríguez, A.C., Solomon, D., González, P., Kreimer, A.R., Porras, C., Schussler, J., Jiménez, S., Sherman, M.E., Quint, W., Schiller, J.T., Lowy, D.R., Schiffman, M., Hildesheim, A., Costa Rica Vaccine Trial Group: Prevention of persistent human papillomavirus infection by an HPV16/18 vaccine: a community-based randomized clinical trial in Guanacaste, Costa Rica. Cancer Discov **1**(5), 408–19 (2011). doi:10.1158/2159-8290.CD-11-0131

10. Stanley, M.: Immune responses to human papillomavirus. Vaccine **24 Suppl 1**, 16–22 (2006). doi:10.1016/j.vaccine.2005.09.002

11. Ferenczy, A., Franco, E.: Persistent human papillomavirus infection and cervical neoplasia. Lancet Oncol **3**(1), 11–6 (2002)

12. zur Hausen, H.: Review: Papillomaviruses — to Vaccination and Beyond. Biochemistry **73**(5), 498–503 (2008). doi:10.1134/S0006297908050027

13. Einstein, M.H., Schiller, J.T., Viscidi, R.P., Strickler, H.D., Coursaget, P., Tan, T., Halsey, N., Jenkins, D.: Clinician's guide to human papillomavirus immunology: knowns and unknowns. Lancet Infect Dis **9**(6), 347–56 (2009). doi:10.1016/S1473-3099(09)70108-2

14. Van Hede, D., Langers, I., Delvenne, P., Jacobs, N.: Origin and immunoescape of uterine cervical cancer. Presse Med **43**(12P2), 413–421 (2014). doi:10.1016/j.lpm.2014.09.005

15. Stanley, M.: Immunology of HPV infection. Curr Obstet Gynecol Rep **4**(4), 195–200 (2015). doi:10.1007/s13669-015-0134-y. Accessed 2017-03-20

16. Gao, W., Weng, J., Gao, Y., Chen, X.: Comparison of the vaginal microbiota diversity of women with and without human papillomavirus infection: a cross-sectional study. BMC Infect Dis **13**, 271 (2013). doi:10.1186/1471-2334-13-271

For peer review only

17. Brotman, R.M., Shardell, M.D., Gajer, P., Tracy, J.K., Zenilman, J.M., Ravel, J., Gravitt, P.E.: Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. J Infect Dis **210**(11), 1723–33 (2014). doi:10.1093/infdis/jiu330

18. Koutsky, L.A., Ault, K.A., Wheeler, C.M., Brown, D.R., Barr, E., Alvarez, F.B., Chiacchierini, L.M., Jansen, K.U., Proof of Principle Study Investigators: A controlled trial of a human papillomavirus type 16 vaccine. N Engl J Med **347**(21), 1645–51 (2002). doi:10.1056/NEJMoa020586

19. Riethmuller, D., Jacquard, A.-C., Lacau St Guily, J., Aubin, F., Carcopino, X., Pradat, P., Dahlab, A., Prétet, J.-L.: Potential impact of a nonavalent hpv vaccine on the occurrence of hpv-related diseases in france. BMC Public Health **15**, 453 (2015). doi:10.1186/s12889-015-1779-1

20. Joura, E.A., Giuliano, A.R., Iversen, O.-E., Bouchard, C., Mao, C., Mehlsen, J., Moreira, E.D. Jr, Ngan, Y., Petersen, L.K., Lazcano-Ponce, E., Pitisuttithum, P., Restrepo, J.A., Stuart, G., Woelber, L., Yang, Y.C., Cuzick, J., Garland, S.M., Huh, W., Kjaer, S.K., Bautista, O.M., Chan, I.S.F., Chen, J., Gesser, R., Moeller, E., Ritter, M., Vuocolo, S., Luxembourg, A., Broad Spectrum HPV Vaccine Study: A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med **372**(8), 711–23 (2015). doi:10.1056/NEJMoa1405044

21. Murall, C.L., Bauch, C.T., Day, T.: Could the human papillomavirus vaccines drive virulence evolution? Proc Biol Sci **282**, 20141069 (2015). doi:10.1098/rspb.2014.1069

22. Moscicki, A.-B., Ma, Y., Wibbelsman, C., Darragh, T.M., Powers, A., Farhat, S., Shiboski, S.: Rate of and Risks for Regression of CIN-2 in adolescents and young women. Obstet Gynecol **116**(6), 1373–1380 (2010). doi:10.1097/AOG.0b013e3181fe777f

23. Buck Jr., H.W.: Warts (genital). BMJ Clin Evid **2015**, 1602 (2015)

24. Herrero, R., González, P., Markowitz, L.E.: Present status of human papillomavirus vaccine development and implementation. Lancet Oncol **16**(5), 206–16 (2015). doi:10.1016/S1470-2045(14)70481-4

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

25. Maver, P.J., Poljak, M.: Progress in prophylactic human papillomavirus (HPV) vaccination in 2016: A literature review. Vaccine (2018). doi:10.1016/j.vaccine.2017.07.113

26. Fagot, J.-P., Boutrelle, A., Ricordeau, P., Weill, A., Allemand, H.: HPV vaccination in France: uptake, costs and issues for the National Health Insurance. Vaccine **29**(19), 3610–6 (2011). doi:10.1016/j.vaccine.2011.02.064

27. YAHIA, M.-B.B.H., DERVAUX, B., DUPORT, N., FLORET, D., GAILLOT, J., HEARD, I., JACQUET, A., GOASTER, C.L., LEVY-BRUHL, D., MORER, I., du CHATELET, I.P., PEIGUE-LAFEUILLE, H., RUMEAU-PICHON, C.: Vaccination contre les infections à papilloamvirus. Technical report, Haut Conseil de la Santé Publique, Paris, France (2014). https://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=454

28. Clarivet, B., Picot, E., Marchandin, H., Tribout, V., Rachedi, N., Schwartzentruber, E., Ledésert, B., Dereure, O., Guillot, B., Picot, M.-C.: Prevalence of Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium in asymptomatic patients under 30 years of age screened in a French sexually transmitted infections clinic. Eur J Dermatol **24**(5), 611–6 (2014). doi:10.1684/ejd.2014.2413

29. Ravel, J., Brotman, R.M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D.W., Sakamoto, J., Koenig, S.S., Fu, L., Zhou, X., Hickey, R.J., Schwebke, J.R., Forney, L.J.: Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. Microbiome **1**(1), 29 (2013). doi:10.1186/2049-2618-1-29

30. Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., Forney, L.J.: Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA **108**, 4680–7 (2011). doi:10.1073/pnas.1002611107

31. Gajer, P., Brotman, R.M., Bai, G., Sakamoto, J., Schütte, U.M.E., Zhong, X., Koenig, S.S.K., Fu, L., Ma, Z.S., Zhou, X., Abdo, Z., Forney, L.J., Ravel, J.: Temporal dynamics of the

For peer review only

human vaginal microbiota. Sci Transl Med **4**(132), 132–52 (2012). doi:10.1126/scitranslmed.3003605

32. Kleter, B., van Doorn, L.-J., ter Schegget, J., Schrauwen, L., van Krimpen, K., Burger, M., ter Harmsel, B., Quint, W.: Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses. Am J Pathol **153**(6), 1731–1739 (1998). doi:10.1016/S0002-9440(10)65688-X

33. Geraets, D.T., Struijk, L., Kleter, B., Molijn, A., van Doorn, L.-J., Quint, W.G.V., Colau, B.: The original SPF10 LiPA25 algorithm is more sensitive and suitable for epidemiologic HPV research than the SPF10 INNO-LiPA Extra. J Virol Meth **215-216**, 22–29 (2015). doi:10.1016/j.jviromet.2015.01.001

34. Gravitt, P.E., Peyton, C.L., Alessi, T.Q., Wheeler, C.M., Coutlée, F., Hildesheim, A., Schiffman, M.H., Scott, D.R., Apple, R.J.: Improved amplification of genital Human Papillomaviruses. J Clin Microbiol **38**(1), 357–361 (2000)

35. Micalessi, I.M., Boulet, G.A.V., Bogers, J.J., Benoy, I.H., Depuydt, C.E.: High-throughput detection, genotyping and quantification of the human papillomavirus using real-time PCR. Clin Chem Lab Med **50**(4), 655–61 (2012). doi:10.1515/cclm.2011.835

36. Hunter, P.J., Sheikh, S., David, A.L., Peebles, D.M., Klein, N.: Cervical leukocytes and spontaneous preterm birth. Journal of Reproductive Immunology **113**, 42–49 (2016). doi:10.1016/j.jri.2015.11.002

37. Shannon, B., Yi, T.J., Perusini, S., Gajer, P., Ma, B., Humphrys, M.S., Thomas-Pavanel, J., Chieza, L., Janakiram, P., Saunders, M., Tharao, W., Huibner, S., Shahabi, K., Ravel, J., Rebbapragada, A., Kaul, R.: Association of HPV infection and clearance with cervicovaginal immunology and the vaginal microbiota. Mucosal Immunology **10**(5), 1310–1319 (2017). doi:10.1038/mi.2016.129

38. Lajoie, J., Juno, J., Burgener, A., Rahman, S., Mogk, K., Wachihi, C., Mwanjewe, J.,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Plummer, F.A., Kimani, J., Ball, T.B., Fowke, K.R.: A distinct cytokine and chemokine profile at the genital mucosa is associated with HIV-1 protection among HIV-exposed seronegative commercial sex workers. Mucosal Immunol **5**(3), 277–287 (2012). doi:10.1038/mi.2012.7

39. Juno, J.A., Boily-Larouche, G., Lajoie, J., Fowke, K.R.: Collection, isolation, and flow cytometric analysis of human endocervical samples. J Vis Exp **89**, 51906 (2014). doi:10.3791/51906

40. Nowak, M.A., May, R.M.: Virus Dynamics: Mathematical Principles of Immunology and Virology. Oxford University Press, Oxford, USA (2000)

41. Stafford, M.A., Corey, L., Cao, Y., Daar, E.S., Ho, D.D., Perelson, A.S.: Modeling plasma virus concentration during primary HIV infection. J. theor. Biol. **203**(3), 285–301 (2000). doi:10.1006/jtbi.2000.1076

42. Perelson, A.S.: Modelling viral and immune system dynamics. Nat. Rev. Immunol. **2**(1), 28–36 (2002). doi:10.1038/nri700

43. Murall, C.L., Jackson, R., Zehbe, I., Boulle, N., Segondy, M., Alizon, S.: Epithelial stratification shapes infection dynamics. bioRxiv, 231985 (2017). doi:10.1101/231985

44. Steimer, J.L., Vozeh, S., Racine Poon, A., Holford, N., O'Neil, R.: The population approach: rationale, methods and applications in clinical pharmacology and drug development. In: Balant, P.G.W..L. (ed.) Handbook of Experimental Pharmacology, vol. 110, pp. 405–451. Springer, Berlin (1994)

45. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. Journal of Statistical Software **67**(1) (2015). doi:10.18637/jss.v067.i01

46. Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., Deng, L., Yeliseyev, V., Delaney, M.L., Liu, Q., Olle, B., Stein, R.R., Honda, K., Bry, L., Gerber, G.K.: MDSINE: Microbial dynamical systems INference engine for microbiome time-series analyses. Genome

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Biology **17**, 121 (2016). doi:10.1186/s13059-016-0980-6. Accessed 2017-03-09

47. Faust, K., Lahti, L., Gonze, D., de Vos, W.M., Raes, J.: Metagenomics meets time series analysis: unraveling microbial community dynamics. Curr Opin Microbiol **25**, 56–66 (2015). doi:10.1016/j.mib.2015.04.004

48. Fox, G.A., Negrete-Yankelevich, S., Sosa, V.J.: Ecological Statistics: Contemporary Theory and Application. Oxford University Press, Oxford, USA (2015)

49. Shi, Y., Li, L., Hu, Z., Li, S., Wang, S., Liu, J., Wu, C., He, L., Zhou, J., Li, Z., Hu, T., Chen, Y., Jia, Y., Wang, S., Wu, L., Cheng, X., Yang, Z., Yang, R., Li, X., Huang, K., Zhang, Q., Zhou, H., Tang, F., Chen, Z., Shen, J., Jiang, J., Ding, H., Xing, H., Zhang, S., Qu, P., Song, X., Lin, Z., Deng, D., Xi, L., Lv, W., Han, X., Tao, G., Yan, L., Han, Z., Li, Z., Miao, X., Pan, S., Shen, Y., Wang, H., Liu, D., Gong, E., Li, Z., Zhou, L., Luan, X., Wang, C., Song, Q., Wu, S., Xu, H., Shen, J., Qiang, F., Ma, G., Liu, L., Chen, X., Liu, J., Wu, J., Shen, Y., Wen, Y., Chu, M., Yu, J., Hu, X., Fan, Y., He, H., Jiang, Y., Lei, Z., Liu, C., Chen, J., Zhang, Y., Yi, C., Chen, S., Li, W., Wang, D., Wang, Z., Di, W., Shen, K., Lin, D., Shen, H., Feng, Y., Xie, X., Ma, D.: A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. Nat Genet **45**(8), 918–22 (2013). doi:10.1038/ng.2687

50. Chen, D., Gaborieau, V., Zhao, Y., Chabrier, A., Wang, H., Waterboer, T., Zaridze, D., Lissowska, J., Rudnai, P., Fabianova, E., Bencko, V., Janout, V., Foretova, L., Mates, I.N., Szeszenia-Dabrowska, N., Boffetta, P., Pawlita, M., Lathrop, M., Gyllensten, U., Brennan, P., McKay, J.D.: A systematic investigation of the contribution of genetic variation within the MHC region to HPV seropositivity. Hum Mol Genet **24**(9), 2681–2688 (2015). doi:10.1093/hmg/ddv015

51. Chen, D., Gyllensten, U.: Lessons and implications from association studies and post-GWAS analyses of cervical cancer. Trends Genet **31**(1), 41–54 (2015). doi:10.1016/j.tig.2014.10.005

For peer review only

52. Winer, R.L., Feng, Q., Hughes, J.P., O'Reilly, S., Kiviat, N.B., Koutsky, L.A.: Risk of female human papillomavirus acquisition associated with first male sex partner. J Infect Dis **197**(2), 279–82 (2008). doi:10.1086/524875

53. Winer, R.L., Hughes, J.P., Feng, Q., O'Reilly, S., Kiviat, N.B., Holmes, K.K., Koutsky, L.A.: Condom use and the risk of genital human papillomavirus infection in young women. N Engl J Med **354**(25), 2645–54 (2006). doi:10.1056/NEJMoa053284

54. Depuydt, C.E., Thys, S., Beert, J., Jonckheere, J., Salembier, G., Bogers, J.J.: Linear viral load increase of a single HPV-type in women with multiple HPV infections predicts progression to cervical cancer. Int J Cancer **139**(9), 2021–2032 (2016). doi:10.1002/ijc.30238

55. WHO HPV LabNet. World Health Organization. http://www.who.int/biologicals/areas/human_papillomavirus/WHO_HPV_LabNet/en/

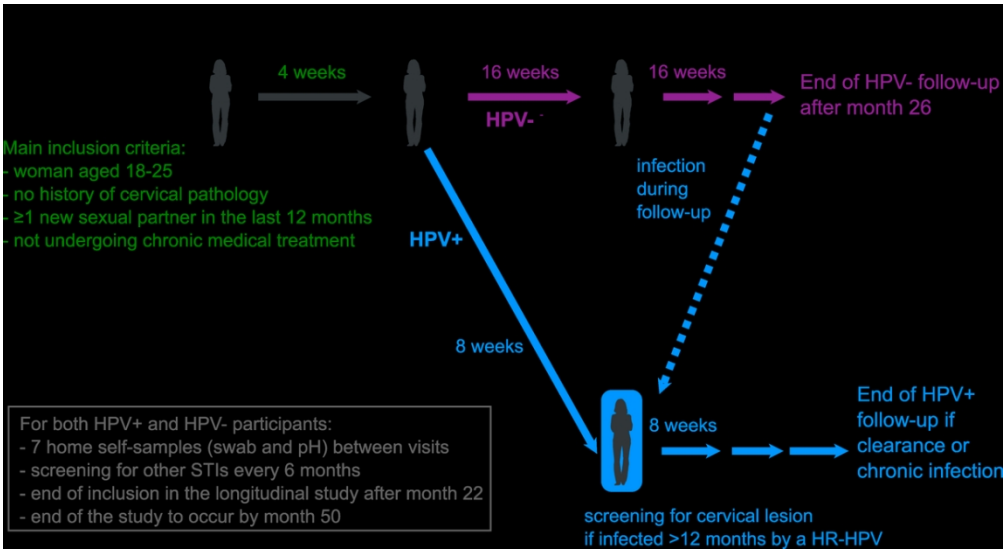
For peer review only

FIGURE CAPTIONS

Figure 1: General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V_1), a results visit (V_2) and then return visits (V_i with $i > 2$). For the cross-sectional study, participants only have V_1 and V_2.

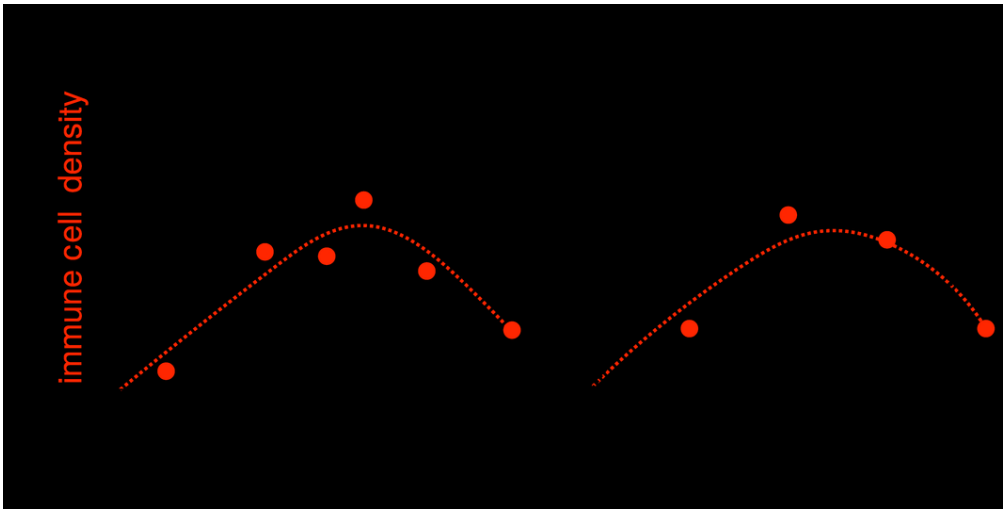
Figure 2: Fitting kinetics dynamical models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

For peer review only



Caption : General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V1), a results visit (V2) and then return visits (Vi with i > 2). For the cross-sectional study, participants only have V1 and V2.

103x56mm (300 x 300 DPI)



Fitting kinetics dynamical models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

90x45mm (300 x 300 DPI)

BMJ Open

The natural history, dynamics, and ecology of Human papillomaviruses (HPVs) in genital infections of young women: the PAPCLEAR study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025129.R1
Article Type:	Protocol
Date Submitted by the Author:	25-Feb-2019
Complete List of Authors:	Murall, Carmen Lia; Centre National de la Recherche Scientifique, MIVEGEC Rahmoun, Massilva; Centre National de la Recherche Scientifique, MIVEGEC Selinger, Christian; Centre National de la Recherche Scientifique, MIVEGEC Baldellou, Monique; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Bernat, Claire; Centre National de la Recherche Scientifique, MIVEGEC Bonneau, Marine; Centre Hospitalier Regional Universitaire de Montpellier, Department of Obstetrics and Gynaecology Boué, Vanina; Centre National de la Recherche Scientifique, MIVEGEC Buisson, Mathilde; Centre Hospitalier Regional Universitaire de Montpellier, Department of Research and Innovation Christophe, Guillaume; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD D'Auria, Giuseppe; Fundacio per al Foment de la Investigacio Sanitaria i Biomedica, Sequencing and Bioinformatics Service; Centro de Investigacion Biomedica en Red de Epidemiologia y Salud Publica De Taroni, Florence; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Foulongne, Vincent; Centre Hospitalier Regional Universitaire de Montpellier, Department of Bacteriology and Virology Froissart, Rémy; Centre National de la Recherche Scientifique, MIVEGEC Graf, Christelle; Centre Hospitalier Regional Universitaire de Montpellier, Department of Obstetrics and Gynaecology Grasset, Sophie; Centre Hospitalier Regional Universitaire de Montpellier, Department of Virology; Centre National de la Recherche Scientifique, MIVEGEC Groc, Soraya; Centre Hospitalier Regional Universitaire de Montpellier, Department of Virology; Centre National de la Recherche Scientifique, MIVEGEC Hirtz, Christophe; Centre Hospitalier Regional Universitaire de Montpellier, LBPC/PPC, IRMB Jaussent, Audrey; Centre Hospitalier Regional Universitaire de Montpellier, Department of Medical Information Lajoie, Julie; University of Manitoba College of Medicine, Department of Medical microbiology Lorcy, Frédérique; Centre Hospitalier Regional Universitaire de

	Montpellier, Laboratoire d'anatomie et cytologie pathologiques Picot, Eric; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD PICOT, Marie-Christine; Centre Hospitalier Regional Universitaire de Montpellier, Department of Medical Information Ravel, Jacques; University of Maryland School of Medicine, Institute for Genome Sciences Reynes, Jacques; Centre Hospitalier Regional Universitaire de Montpellier, Department of Infectious and Tropical Diseases Rousset, Thérèse; Centre Hospitalier Regional Universitaire de Montpellier, Department of pathology and oncobiology Seddiki, Aziza; Centre Hospitalier Regional Universitaire de Montpellier, Department of Research and Innovation Teirlinck, Martine; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Tribout, Vincent; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Tuaillon, Edouard; Centre Hospitalier Regional Universitaire de Montpellier, Department of bacteriology and virology Waterboer, Tim; Deutsches Krebsforschungszentrum, Infections and Cancer Epidemiology Jacobs, Nathalie; Universite de Liege Faculte des Sciences, GIGA-Research, Cellular and molecular immunology Bravo, Ignacio; Centre National de la Recherche Scientifique, MIVEGEC Segondy, Michel; Centre Hospitalier Regional Universitaire de Montpellier, Department of Bacteriology and Virology Boulle, Nathalie; Centre Hospitalier Regional Universitaire de Montpellier, Department of pathology and oncobiology Alizon, Samuel; Centre National de la Recherche Scientifique, MIVEGEC
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology, Immunology (including allergy), Genetics and genomics
Keywords:	VIROLOGY, IMMUNOLOGY, Epidemiology < INFECTIOUS DISEASES, MICROBIOLOGY, GENETICS

SCHOLARONE™
Manuscripts

STUDY PROTOCOL

The natural history, dynamics, and ecology of Human papillomaviruses (HPVs) in genital infections of young women: the PAPCLEAR study

Carmen Lía Murall[1] , Massilva Rahmoun[1], Christian Selinger[1], Monique Baldellou[2], Claire Bernat[1], Marine Bonneau[3] , Vanina Boué[1] , Mathilde Buisson[4], Guillaume Christophe[2], Giuseppe D'Auria[5,6], Florence De Taroni[2], Vincent Foulongne[7,8], Rémy Froissart[1], Christelle Graf[3], Sophie Grasset[1,2], Soraya Groc[1,7], Christophe Hirtz[9], Audrey Jaussent[10] , Julie Lajoie[11], Frédérique Lorcy[12], Eric Picot[2], Marie-Christine Picot[10], Jacques Ravel[13], Jacques Reynes[11], Thérèse Rousset[12], Aziza Seddiki[4], Martine Teirlinck[2], Vincent Tribout[2], Édouard Tuaillon[8], Tim Waterboer[15] , Nathalie Jacobs[16], Ignacio G Bravo[1], Michel Segondy[7,8], Nathalie Boulle[8,12] and Samuel Alizon[1,*]

For peer review only

Word count: 5105 words, excluding title page, abstract, references, figures and tables.

1 Laboratoire MIVEGEC (UMR 5290 CNRS, IRD, UM), 911, avenue Agropolis, BP 64501, 34394 Montpellier, France
2 Center for Free Information, Screening and Diagnosis (CeGIDD), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
3 Department of Obstetrics and Gynaecology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France.
4 Department of Research and Innovation (DRI), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
5 Sequencing and Bioinformatics Service, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública), 46020 Valencia, Spain
6 CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain
7 Department of Virology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
8 Pathogenesis and Control of Chronic Infections, INSERM, CHU, University of Montpellier, Montpellier, France
9 University of Montpellier, LBPC/PPC- IRMB, CHU de Montpellier, 80 rue Augustin Fliche, Montpellier, France
10 Department of Medical Information (DIM), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
11 Department of Medical microbiology, University of Manitoba, 745 Bannatyne, Winnipeg, Canada
12 Department of pathology and oncobiology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
13 Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, USA
14 Department of Infectious and Tropical Diseases, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
15 German Cancer Research Center (DKFZ), Infections and Cancer Epidemiology, Im Neuenheimer Feld 280, Heidelberg, Germany
16 GIGA-Research, Cellular and molecular immunology, University of Liège, 3 Avenue de l'Hôpital, 4000 Liège, Belgium

* Author for correspondence: samuel.alizon@cnrs.fr

For peer review only

Abstract

Introduction

Human papillomaviruses (HPVs) are responsible for one third of all cancers caused by infections. Most HPV studies focus on chronic infections and cancers, and thus, we know little about the early stages of viral infection. In particular, the effects of the dynamic interactions between the immune system, the microbiota, and the viral and host genetics on infection clearance or persistence remains poorly understood.

Methods and Analysis

We follow 150 women, aged 18-25 years, longitudinally to monitor immune response features (cytokines and immune cells in the genital tract, circulating anti-HPV antibodies), virus load of HPVs, and vaginal microbiota composition. This is complemented by the assessment of viral and human genetics and behavioural data. To increase the statistical power of the epidemiological arm of the study, an additional 150 women are screened cross-sectionally. This study will provide one of the most detailed follow-up studies of acute HPV infections and their interactions with the host and the vaginal microbiota. It will also allow us to investigate related issues regarding HPV intra-host evolution and diversity, vaginal microbiota dynamics, and sexually transmitted infections.

Ethics and Dissemination

This study has been approved by the Comité de Protection des Personnes Sud Méditerranée I (reference number 2016-A00712-49); by the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (reference number 16.504); by the Commission Nationale Informatique et Libertés (reference number MMS/ABD/AR1612278, decision number DR-2016-488) and by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (reference 20160072000007). The results will be published in preprint servers, peer-reviewed journals and disseminated through conferences.

Trial registration number: NCT02946346

Keywords: HPV; acute infection; persistence; virus load; immunity; microbiota; viral kinetics

For peer review only

Article summary

Strengths and limitations of this study

- Dense follow-up (visit every two months for infected women with additional self-sampling every week) for N=150 women.

- Combination of virological (virus load), immunological (cytokine concentrations and immune cell percentages) and environmental (vaginal microbiota composition, pH) measurement at each visit.

- A limitation is that the density of the follow-up limits the number of participants, which can affect analyses at the epidemiological level.

- We complement the longitudinal study with a cross-sectional study of N=150 women to allow for epidemiological analyses.

For peer review only

Introduction

Epidemiology of HPV genital infections in young adults and public health implications

Infections by Human Papillomaviruses (HPVs) are likely the most common sexually transmitted infection (STI) globally. It is often estimated that, worldwide, 75% of the individuals will be infected at some point in their life by an HPV type [1]. In France, a study performed in 2013 in the Paris area estimated the prevalence of genital infections by HPVs at ≈ 25% of women below 25 years of age [2]. In the area of Montpellier, prevalence of oncogenic HPVs (often referrer to as 'high-risk', HR) in pregnant women aged 16 to 42 years was close to 20% [3]. These numbers are consistent with worldwide estimates according to which HPVs are most prevalent in women under 25 years of age, with an estimated overall prevalence of 24% [4].

Fortunately, the vast majority of infections by HPVs are asymptomatic and benign. Even for HPV16, which is probably the most oncogenic biologic agent to humans, only a minority of infections (less than 10%) become persistent [5], and then a minority of these (12%) progress to cancer if untreated [1, 6]. Indeed, it is estimated that approximately 70 to 100% of infections by HPVs are cleared within 12 to 24 months, with strong differences between virus types [5, 7–9]. Recent studies suggest that primo-infections could be shorter in young girls [10] but, in general, there are many unknowns about the biology of non-persisting infections [11].

Our lack of knowledge partly comes from the fact that in vaccine trials, from which most of the data on infection duration originate, participants are followed every six months for several years [5, 7, 9, 12]. This frequency is sufficient to estimate the time to clearance (or to persistence) but it is not precise enough to understand the within-host dynamics, often referred to as 'kinetics' [13], of infections that last on average 6 to 24 months. Arbitrarily, after 24 months of infection, an infection is often considered as being persistent [14].

Some factors have been shown to correlate with persistence (e.g. immunosuppression, smoking, and co-infection with other STIs [15]) but we do not know how these affect viral

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

kinetics. Also, some changes in viral-immunity interactions appear to be related to persistence and disease progression [16–19] but, again, we do not know the underlying interactions between the viruses, the host target cells, and the immune response in acute infections [11]. Finally, it has been argued that the vaginal microbiota may differ between HPV-infected and HPV-uninfected women [20] and that specific microbiota composition may interact with HPV detection [21]. However, it is difficult to disentangle the cause and the consequence. For instance, does the microbiota composition change after the establishment of an HPV infection, or do certain microbiota compositions increase susceptibility to HPV infection?

A better understanding of the within-host infection dynamics and of the determinants of clearance and persistence of viral infection is particularly important in the context of vaccination [22–25]. Indeed, the long-term efficacy of the anti-HPVs vaccines at the population level will largely depend on the within-host viral dynamics because, ultimately, most selective pressures on viral populations occur via the immune response [26]. Furthermore, a better understanding of acute HPV infections can shed a new light on issues related to latency, fertility, or immunotherapies [11].

Prevention strategies and treatment

Treatment

Since most infections by HPVs are benign in young adults and clear within six to 24 months, the current standard of care is to avoid over-treatment, even in the presence of cervical lesions [27]. Clinical interventions (colposcopies, biopsies, and surgery) are less often performed with young women (< 25 years) and only for high-grade (pre-cancerous) lesions (cervical intraepithelial neoplasia grade 2, CIN-2, or more). Low-grade lesions (CIN-1) are not systematically treated but rather monitored yearly to detect any progression to high-grade lesions.

Genital warts caused by non-oncogenic HPVs (often referred to as ‘low-risk’, LR, HPVs) can be removed by surgery or treated with bi- and trichloroacetic acid, cryotherapy or other treatments [28].

For peer review only

HPV vaccination

There are currently three licensed vaccines: a bivalent vaccine (Cervarix®) targeting HPV16 and HPV18 (together accounting for 70% of cervical cancers [1]), a quadrivalent vaccine (Gardasil®) that additionally targets HPV6 and HPV11 (non-oncogenic, but highly prevalent and associated to benign proliferative lesions) and, since 2014, a nonavalent vaccine (Gardasil 9®) that targets five more oncogenic types (HPV31, HPV33, HPV45, HPV52, and HPV58, which altogether account for 20% of cervical cancers [24]). These vaccines succeed in eliciting a protective immune response against new infections by the targeted viruses, and are used throughout the world, albeit with wide variation in coverage (for reviews, see e.g. [29, 30]).

Vaccination campaigns in France started in 2006 but with limited coverage: it reached 28.5% in 2008 [31] and has been decreasing ever since [32]. The vaccine is recommended for girls from 11 to 14 years of age, currently with a vaccination scheme of two doses with a six months interval. A catch-up is organised for girls aged 15-19 years, with a three-doses vaccination scheme. Vaccination is reimbursed by the French Social Security but is not mandatory. It is also recommended for men who have sex with men (MSM) as well as for immuno-compromised people [32]. Vaccination is now the primary prevention strategy against cervical cancers.

Screening

In France, the secondary prevention strategy against cervical cancer is routine individual cytology-based screening for pre-cancerous and cancerous cervical lesions in women between 25 and 65 years. Cytology can also be performed in younger women if they report risk factors for cervical cancer (multiple partners, chronic STIs or HIV infection [32]). Detection of oncogenic HPVs is proposed for triage in case of abnormal cytology (i.e. high-grade or low-grade squamous intraepithelial lesion, HSIL and LSIL respectively, or Atypical Squamous Cells of Undetermined Significance, ASCUS).

Primary objectives

The first primary objective is to decipher the kinetics and ecology of cervical HPV infections in healthy young women, i.e. follow the population dynamics of the virus, the target epithelial cells,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

and the immune effectors.

The second primary objective is to characterise the diversity of genital HPVs in young women in the region of Montpellier in relationship with their lifestyle, vaccination status, vaginal microbiota, and human genetics.

Secondary objectives

A secondary objective is to characterise the acquisition and clearance dynamics of cervical HPV infections as a function of viral diversity, host immunity, vaginal microbiota and human genetics.

A final objective is to investigate variations in genetic diversity of HPVs during cervical infections.

Methods and analysis

Participants

The study population is composed of young women at risk of HPV infection. The age class was chosen because the prevalence of HPV is the highest (24% worldwide [4] and is approximately 25% in France [2]). Inclusion of younger women would have raised technical issues because of the requirement for parental consent.

Women are recruited through a social media page, and through posters and leaflets distributed at the Universities in Montpellier and at the Montpellier STI screening centre (Centre Gratuit d'Information de Dépistage et de Diagnostic, CeGIDD). The composition of the population visiting the CeGIDD has already been documented in an earlier study [33]. In total, the centre is visited by approximately 3,000 women per year, the majority of which are under 25 years of age (80%). Approximately 40% of the attendants report three or more partners over the last twelve months and approximately 50% report using adequate behaviour for prevention against HIV.

For peer review only

Inclusion criteria Participants are women from 18 to 25 years old living in the metropole of Montpellier. They must be sexually active with at least one new partner over the last 12 months. This criteria is fixed to maximise the incidence of new HPV infections. As in any clinical study, participants must be able to and willing to give written informed consent: they must sign an informed consent form, understand the requirements for the study, and be affiliated to a French social security scheme (which is a state requirement).

Women cannot be included in the study if they have a history of HPV-associated pathology (genital warts or cervical lesions), if they are pregnant or intending to become pregnant in the coming year, infected by HIV, undergoing (or planning to undergo) intense medical treatment (biotherapy, chemotherapy, immunosuppression), planning on moving outside the Montpellier metropolitan area within the next 18 months, in a dependency or employment with the sponsor or the investigator, if they participated in a clinical trial involving administration of drugs within the last four weeks or if they belong to a vulnerable group (e.g. children, adults with physical or mental disabilities).

Design/setting

This study has a longitudinal component aimed at deciphering within-host dynamics and a cross-sectional component, aimed at understanding the diversity of HPV infections in young adults in the area of Montpellier, France. The general structure of the study is shown in Fig 1.

If a woman fits the main inclusion criteria, she can go through an inclusion visit (V_1) with a physician (gynaecologist or midwife) at the CeGIDD. During this visit, she presents the study and checks all inclusion criteria before asking the participant to read and sign the informed consent form. Participants then undergo a medical consultation during which a number of samples are collected (see below). They then fill out health and lifestyle questionnaires and are given cotton-flocked swabs for self-sampling at home the next visit, along with instructions on how to fill in weekly questionnaires through an online form (these are performed throughout the study).

For peer review only

An appointment is scheduled four weeks later for the Results visit (V_2), where the cervical cytology results are communicated. We collect some samples and provide more self-sample swabs for home collection.

The next return visits (V_i, where $i > 2$) are as follows:

- Participants with a positive DEIA HPV test (see below), i.e. infected by an *Alphapapillomavirus*, at V_1 join the HPV positive (HPV+) arm of the study with return visits scheduled every 2 months.
- Participants with a negative DEIA HPV test at V_1 join the HPV negative (HPV-) arm with return visits scheduled every 4 months.
- HPV- participants infected by an *Alphapapillomavirus* move to the HPV+ arm.

Intervals between visits are based on earlier results showing that HPV infections last from 9 to 18 months on average depending on the HPV type [5, 7–9] and that a follow-up of 4 months yields results that are difficult to analyse [21]. The longer interval in the HPV- arm is based on the estimated incidence for HPV genital infections in young women, which is greater than 30% [34, 35].

Participants in the HPV- arm are followed until month 32 of the study.

Participants in the HPV+ arm are followed until they clear the infection or until they have been infected for 24 months (after which we consider that the infection is persistent). Clearance is defined as being negative at two visits in a row for the first HPV type detected in the follow-up.

In between these visits to the CeGIDD, participants are asked to perform regular (every week for HPV+ and every second week for HPV-) self-samples using vaginal swabs, along with a measure of vaginal pH and filling a short questionnaire. Self-samples are stored in the participants' freezer and brought back at every visit.

The study will end with the last HPV+ participant having cleared the infection or been infected for 24 months.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Patients and public involvement

Since all participants are healthy, they are referred to as participants rather than patients. As in any longitudinal study, ensuring participant commitment is challenging. To achieve this goal, we have set up a compensation of 40 EUR per visit and an additional 10 EUR in case of a complete follow-up. Furthermore, participants who have answered a sufficient number of questionnaires and brought back a sufficient number of self samples will get a 100 EUR bonus at the end. Overall, a participant performing 12 return visits would gain a total compensation of 650 EUR.

Participants did not play a role in the design of this study.

Results of the study will be disseminated to participants who have left the study and to the general public via an email newsletter in French.

Visits

The summary of the visit schedule and of the samples collected at each visit is shown in Table 1.

Inclusion visit (*V*1)

This visit takes place at the CeGIDD and is scheduled by the Clinical Research Technician (TEC) via phone or email.

Women meet a study investigator, who explains the goals and requirements of the study. The physician also checks that the inclusion criteria are met. If so, after a general discussion, the informed consent forms are signed.

The female physician/midwife performs a general exam and then a gynaecological exam during which the following samples are taken:

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

- vaginal pH cotton swab (EcoCareTM),
- vaginal swab (Copan ESwabTM) in 1mL Amies liquid for DNA extraction and microbiota analysis,
- vaginal swab (Copan ESwabTM) in 1mL of RNA preservation medium,
- ophthalmic sponge (Weck-cel®) to collect cervical secretions for cytokines analysis,
- cervical smear in 20mL of Thinprep® (Preservcyt® liquid) for HPV and HSV assays, and cytology evaluation.

Following the gynaecological consultation, the participant meets with a nurse to measure body temperature, blood pressure and draw 20mL of blood (a 5mL tube for SNPs sequencing, a 10mL tube for immunophenotyping and a 5mL tube for HPV antibody titration). For the longitudinal study, the nurse provides the participant with 3 self-sampling kits, 3 pH strips, a freezer box to bring back to the next visit, as well as instructions on how to perform the home sampling and store the samples in her personal freezer until the next visit.

If the participant has not been tested for a STI in the last 3 months, the nurse draws an additional blood tube of 5mL to test for STIs (HIV, HCV, HBV) and collects vaginal self-samples for chlamydiae and gonorrhea detection. Syphilis testing is prescribed to participants who meet the STI clinic's guidelines.

Finally, the participant meets with the TEC to fill in questionnaires #1 (inclusion visit) and #3 (home). The TEC answers any remaining questions, explains how to fill the home questionnaires (#3) and sets an appointment for the Results visit.

Results visit (*V*2)

During this visit, the participants are given the result of cervical lesion screening using the liquid cytology (normal, ASCUS, LSIL or HSIL). Participants with a high-grade lesion (HSIL) exit the study and are referred to the gynaecology service of the CHU of Montpellier.

During this visit, the physician/midwife collects additional samples: 2 vaginal swabs for DNA and RNA analysis, and a cervical smear in 10mL of PBS (to confirm HPV status and perform flow

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

cytometry analyses).

The participant fills in questionnaires #2 (for return visits) and #3 (home). An appointment for the next visit is set and swabs for home self-sampling are given.

Return visits (*Vi*)

These visits only occur in the longitudinal study.

HPV- arm Participants uninfected by HPV visit the clinic every 4 months until month 26. During these visits, the same samples as in the inclusion visit (V_1**)** are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

The nurse only draws blood if a screening test for STIs other than HPV is required. The participant then fills in questionnaires #2 and #3 and an appointment is set for the next visit in 16 weeks.

If an HPV infection is detected in the cervical smear collected during this visit, the participant moves to the HPV+ arm and the TEC contacts the participant to move her appointment forward.

HPV+ arm Participants infected by HPV visit the clinic every 2 months. They cannot switch arm and will remain in the HPV+ arm until clearance or the end of the study. During the visits, the same samples as in the inclusion visit (V_0) are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

The nurse then draws 5mL of blood for HPV antibody titration. If this is the first HPV+ visit following an HPV- visit, the nurse also draws 10mL of blood for immunophenotyping. Finally, if a test for additional STIs is needed, the nurse draws 5mL of blood and collects vaginal self-samples for STI detection.

Importantly, if the participant has been infected by a HR-HPV for more than 12 months and cytology has not been performed within the last 12 months, the cervical smear is put in Thinprep® fixation medium, instead of PBS, for cytological analysis (cervical lesion screening).

For peer review only

Finally, the participant fills in questionnaires #2 and #3, receives self-samples for home collection and an appointment is set for the next visit in 8 weeks.

Endpoints

The primary endpoint for the study is the inclusion and follow-up of HPV-infected women in order to describe the kinetics of HPV virus load, and the associated immune response.

Secondary endpoints are the characterisation of the interactions between the course of the infection (e.g. duration), the HPV type(s), the abundance and taxonomic diversity of bacteria, fungi and viruses in the vaginal microbiota, human genetics (SNPs) and basal immunological status.

Technical procedures

DNA extraction

DNA extraction from cervical smears will be performed using Nuclisens EasyMAg from Biomerieux or an equivalent protocol. For the microbiota analyses, special kits involving physical (via beads) and/or enzymatic breaking of the cellular barrier will be favoured following standard protocols to study the vaginal microbiome [36], e.g. the MagAttract® PowerMicrobiome® DNA/RNA kit from Qiagen.

HPV detection, typing and quantification

The participants' infection status (HPV+ or HPV-) will be assessed using the DEIA test, which is based on a PCR of the short SPF10 amplicon [37] and detects all *Alphapapillomaviruses* with great sensitivity.

If the DEIA test is positive, HPVs will be typed using the LiPA$_{25}$ kit, which is based on the same SPF10-PCR, and has a lower detection threshold compared to other hybridisation-based typing methods [38].

The reason for basing the detection on the DEIA rather than the LiPA$_{25}$ is that some *Alphapapillomavirus* may be detected by DEIA but not genotyped by LiPA and also that the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

DEIA is more sensitive than the LiPA. If the DEIA is positive and the LiPA$_{25}$ is negative, typing will be performed by sequencing the product of a PGMY09/11 PCR [39], which targets another region of the HPV genome than the SPF10 PCR.

The quantification of HPV DNA genome copy number in the samples will be performed using the protocol set up by Micalessi et al. [40].

Cytokine titration

Cytokines can be used as markers of immune activation or immunosuppression and can also inform us on which components of the immune system are involved. Cervical sponges are centrifuged after the addition of 175μL of PBS. Cervical secretions are analysed for a set of 5 to 6 cytokines levels using the Meso Scale Discovery (MSD) Multiplex ELISA platform, which has a low detection threshold and a slowly saturating dose-response curve. Based on earlier results [41, 42], we will first investigate a large panel of 20 cyctokines (IFN-α2a, IFN-γ, IL-1α, IL-5, IL-6, IL-8, IL-10, IL-12, IL-15, IL-17, IL-18, IL-23, IL-25, IP-10, MCP-1, MIP-1α, MIP-3α, MIP-3β, TNF-α, TNF-β) to choose the most relevant ones for a longitudinal follow-up.

Flow cytometry

Analysing immune cells via flow cytometry is extremely challenging on cells as fragile as the ones from cervical smears. However, several studies suggest that this is feasible [41–43]. Here, we follow the protocol described in [44].

Stainings are performed using a Duraclone custom mix targeting CD45, CD3, CD4, CD8, CD16, CD56, CD69, CD161 and TCR$\gamma\delta$. The last marker, Live&Dead tests for cellular viability. Samples are acquired using a Navios flow cytometer (Beckman Coulter, three-laser configuration).

Sequencing

Sequencing will be performed for microbiota profiling. It involves PCR amplification of the V3-V4 region of 16S RNA for bacteria [45] and ITS1 for fungi [46]. We anticipate that the bacteria should belong to the operational taxonomic units (OTU) described in the five community state

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

types found in vaginal communities [47, 48]. The virome will also be explored using shotgun sequencing and rolling circle PCR amplification [49]. Human genetics are explored using chip sequencing for SNPs.

Statistical analyses

Times series analyses

The core results of the study will come from the longitudinal follow-up of infected women, which will generate time series, i.e. a set of values collected from the same individual over time (Figure 2). There will be several time series per individual (virus load, number of immune cells, cytokine and antibody levels). These time series will be used to fit mathematical viral kinetics models that describe the interaction between viruses, host target cells (here, in the case of HPV, keratinocytes) and the immune response. These models are commonly developed for viral infections [13, 50–52], including those caused by HPVs [53]. We anticipate our follow-up to yield adequate data for such a fit based on the estimated duration of HPV infections (9 to 18 months [5, 7–9]). Furthermore, the weekly self-samples allow us to increase the resolution if necessary.

We will use non-linear mixed effect models [54] to jointly analyse time series from all participants. More precisely, we will rely on *R* packages such as nlme [55] or lme4 [55]. Note that, in addition to estimating model parameters (e.g. life-expectancy of infected cells or virion production rate of infected cells), this approach also allows us to compare biological models using statistical tools based on model likelihood such as Akaike Information Criterion. For an example of such analysis in the case of HIV, see [51].

Microbiota dynamics

The composition of the vaginal microbiota has already been described and shown to exhibit considerably less diversity than the gut microbiota [47]. The dynamics of this microbiota has also been studied and shown to closely follow menstrual cycles [48].

Using the time series of OTU abundances (measured via 16S RNA sequencing and qPCR) we

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

will infer interaction parameters by assuming an underlying Lotka-Volterra competition model [56]. This work will include time series analysis techniques (e.g. auto-correlation or local similarity analysis) and statistical inference methods in order to infer community structure and interactions from the next-generation sequencing (NGS) datasets [57]. Finally, statistical methods from ecology will also be used to study community diversity (e.g diversity indices) and community assembly, such as cluster and ordination analyses [58].

Genome Wide Association Studies

We will use human single nucleotide polymorphisms (SNPs) inferred by chip sequencing to look for genetic determinants of key traits (e.g. microbiota composition or HPV infection duration). This is classically done by performing a Genome Wide Association Study (GWAS), which is a complex regression method designed for situations where there are many explanatory variables (here millions of SNPs) for a single trait of interest. GWAS will be performed using classical methods [59]. Earlier GWAS studies have been applied to HPV infections for instance to test for determinants to the ability to seroconvert following infection [60] and cervical cancer (see [61] for a review). Here, our expected sample (N = 300 women) is limited but SNPs with large effects have been detected by studies with comparable sizes [62].

Additional analyses

For all collected variables, descriptive statistics will be calculated according to the level of measurement. For metric variables these statistics can be mean and standard deviation as well as quantiles and more robust statistics [63]. In case of categorical variables group proportions and contingency tables are prepared.

Univariate inferential statistics follow a descriptive analysis. Generally, parametric testing procedures are preferred to non-parametric tests, as the former have higher power. That is why, for metric variables, we will first check whether the data can be assumed to be normally distributed. For normally distributed variables, ANOVA statistics are done to detect differences between groups. In case of significance, post-hoc analysis (Tukey test) are planned to reveal pairwise differences. If the data are not normally distributed or ordinally scaled, non-parametric

For peer review only

analyses will be used. These contain the Kruskal-Wallis test and the Wilcoxon test as a post-hoc test with an appropriate correction of the significance level. Since the cell counts are expected to be small, Fisher's exact test will be performed for contingency tables instead of the asymptotic χ^2 test for categorical variables.

Sample size calculation

The study will enrol a total of N = 300 women, with N = 150 in a longitudinal study and N = 150 in a cross-sectional study. The goal of the longitudinal study is to follow 75 women longitudinally, preferentially before they are infected (see above). For the following calculations, we assumed a high percentage of lost during follow-up (30%).

With 150 enrolments and considering that the prevalence of HPV infection in young women is ≈ 60% (based on our preliminary data) and 30% of lost to follow-up, we expect to detect (and successfully follow) 63 infections at inclusion [CI95: 51 − 75, assuming a binomial distribution to calculate the 95% confidence interval].

Among women who are uninfected at the first visit and considering the yearly incidence being close to 30% [64], we expect 12 [CI95: 6 − 20] to be infected during the first year of follow-up (still assuming 30% of lost to follow-up).

In the end, with 150 enrolments and assuming a high percentage of lost to follow-up (30%), we expect to successfully follow 75 [CI95: 56 − 95] women infected at different stages of HPV infection: beginning, during and end.

This will be made possible by the probability of transmission of HPV, which is estimated to be ≈ 90% without condom use and still high with condom use (≈ 40%) [34].

Finally, regarding potential interference with the HPV vaccines, we do not anticipate any

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

significant problem for two reasons. First, as mentioned above, the vaccine coverage is low in France [32]. Second, and more importantly, the vaccines only target few HPV types, thus leaving open the possibility of infection by dozens of types. Furthermore, studying the kinetics of a non-vaccine HPV type in a vaccinated woman will be extremely informative, e.g. to detect any potential cross-reactivity [65].

To run cross-sectional analyses (especially on the microbiota and human genetics), we will enrol *N* = 150 additional women who will only perform the inclusion and the results visits. This sample size was chosen to reach that of earlier GWAS studies [61, 62].

Trial governance

Sponsor

This study is sponsored by the Centre Hospitalier Universitaire (CHU) of Montpellier. The CHU is involved in the implementation of the trial, legal/ethical submissions (see below for details on Ethics approval) and implementing the clinical database (eCRF), which is hosted by Ennov-Clinical (ClinSight). The CHU is not involved in the analysis or interpretation of the data. The CHU of Montpellier performs regular quality control assessments. A clinical research assistant will visit the CeGIDD every 4 months to ensure that implementation is in accordance with the protocol. The CHU has taken out insurance from the Société hospitalière d'assurances mutuelles, 18, rue Edouard Rochet-6 9372 Lyon cedex 08 (contract number 138983) through the full research period, covering its own civil liability and that of any agent (clinical or research staff), in accordance with article L.1121-10 of the French Public Health Code.

Scientific committee

The scientific committee comprises the study investigators, clinicians, scientific experts and representatives of the sponsor. The committee meets yearly and is responsible for following research progress, monitoring compliance with good clinical practices and patient safety. It can also decide relevant modification of the protocol. Requests from third parties to access data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

collected during the study will be evaluated by the committee.

Monitoring

Monitoring is performed during the whole study at CeGIDD according to the sponsor specific SOP. Routine monitoring visits are made by the monitors designated by the sponsor to check compliance with the protocol, the completeness, accuracy and consistency of the data, and adherence to GCP. The principal investigator ensures that eCRFs are completed in a timely manner and must allow periodical access to eCRFs, patient records, drug logs, and all other study-related documents and materials. The frequency of monitoring visits is determined by factors such as study design and the site enrolment requirements but visits will normally occur at least once every 4 months.

Trial registration

The trial has been registered to ClinicalTrials.gov on 27 Oct 2016 with ID number NCT02946346.

Discussion

Expected results

Acute infections by HPVs are important to study because vaccination is most effective when performed before the first infection. However, we currently know very little about the early stages of HPV infections. This clinical study will give us an unprecedented level of detail into the natural history of HPV infections in young women. Variations in virus load over time have been studied but in the context of cervical cancer in older women [66]. In addition, we will also describe the nature and the dynamics of the immune response (local immune cells and cytokines, circulating anti-HPV antibodies) and of the vaginal microbiota. Beyond these kinetics, we will also have access to data such as infection clearance or not in 24 months, presence of more than one HPV type or coinfection by other STIs.

For peer review only

These data will be analysed in the light of numerous cofactors. One of the most important will be human genetics, with the sequencing of millions of SNPs. Others will be related to the sexual behaviour (number of partners, contraception methods, sexual practices) and general lifestyle. We, therefore, expect broader insights regarding sexual health in young women.

Practical and operational issues

One of the main practical challenges resides in the analysis of cervical smears by flow cytometry. Indeed, the tissues are known to be fragile, adhesive and auto-fluorescent. Even though standard protocols now exist [44], they require processing fresh samples in less than 2 hours.

Another potential issue has to do with contamination by HPV DNA between samples, which are frequent in the HPV field due to the robustness of the virions and the sensitivity of the tests. To certify our ability to control for these, we have entered the 2017 GLOBAL HPV DNA Proficiency Panel from the WHO HPV LabNet [67].

Regarding the enrolment of the participants, we do not expect issues with enrolling 150 women in 28 months for the longitudinal study and 150 for the cross-sectional study. This is due to the number of visitors of the centre who fit the inclusion criteria (more than 3,000 per year) and because of earlier high participation rates in the same population ([33] enrolled 1381 participants in 5 months for their study).

Concerning the follow-up, the high incidence rate of HPV can also lead to transient carriage, i.e. women who are positive for a type only at a single visit. This has been observed for instance in longitudinal studies with a tight follow-up interval [21]. To control for this, we will run the HPV detection test on the cells from the cervical smear after washing with RPMI.

For peer review only

Table 1: Summary of the visit schedules and samples take. The cross-sectional study only includes the first two columns (V1 and V2). The ¤indicate samples taken at visits. + participants infected by a HR-HPV for 12 month will have one PBS smear replaced by a Thinprep® smear to perform a cytology and check for lesions. ⊲ this sample is only taken at the first HPV+ visit of a formerly HPV- participant. ⋆ STI detection will be performed at inclusion unless the participant has been tested within the last 3 months and during the study every 6 months if a new partner has been reported or upon request.

For peer review only

	Inclusion ($V1$)	**Results ($V2$)**	**Return (Vi, with $i > 2$)**	
Participants	all	all	HPV+	HPV-
Time	day 0	+ 4 weeks	+ 8 weeks	+ 16 weeks
Eligibility	¤			
Consent	¤			
Gynecological consult	¤	¤	¤	¤
Vaginal pH coton swab	¤	¤	¤	¤
2 vaginal swab samples (Copan ESwabTM)	¤	¤	¤	¤
1 ophtalmological sponge sample	¤		¤	¤
1 cervical smear in Thinprep® (cytology)	¤		+	
1 cervical smear in PBS		¤	+	¤
Blood sampling (HPV antibodies)	¤		¤	
Blood sampling (sequencing)	¤			
Blood sampling (immunophenotyping)	¤		◁	
Other STI detection	⋆	⋆	⋆	⋆
Questionnaire #1 (inclusion)	¤			
Questionnaire #2 (visit)		¤	¤	¤
Questionnaire #3 (home)	¤	¤	¤	¤
Returning self-sampling samples		¤	¤	¤
Serious Adverse Event collection		¤	¤	¤

Abbreviations

ANOVA: Analysis of variance,

ASC-US: Atypical squamous cells of undetermined significance,

CD: Cluster of differentiation,

CI95: 95% Confidence interval,

CeGIDD: Centre Gratuit d'Information de Dépistage et de Diagnostic,

CHU: Centre Hospitalier Universitaire,

CIN: Cervical intraepithelial Neoplasia,

ELISA: enzyme-linked immunosorbent assay,

GWAS: Genome Wide Association Study,

HIV: Human Immunodeficiency Virus,

HPV: Human Papillomavirus,

HR: high-risk,

ITS: Internal Transcribed Spacer,

HSIL: High grade Squamous Intraepithelial Lesion,

LR: low-risk,

LSIL: Low grade Squamous Intraepithelial Lesion,

NGS: Next Generation Sequencing,

OTU: Operational Taxonomic Unit,

PBMC: Peripheric Blood Mononuclear Cell,

PBS: Phosphate Buffered Saline,

RPMI: Roswell Park Memorial Institute medium,

SNP: Single Nucleotide Polymorphism,

TCR: T-cell receptor,

WHO: World Health Organisation.

For peer review only

Trial status

The study began on Oct 1, 2016 and the first inclusion was on Nov 3, 2016. On Jun 23, 2018, 89 participants have been included in the longitudinal study. Inclusions in the longitudinal study will continue until March 2019 and the study is expected to last until Aug 2021.

Conflicts of interests

The authors have read and understood BMJ policy on declaration of interests and declare that they have no competing interests.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 648963). The authors acknowledge further support from the Centre National de la Recherche Scientifique (CNRS), the Institute de Recherche pour le Développement (IRD) and the Centre Hospitalier Universitaire (CHU) of Montpellier.

Data statement

All personal and identifying information collected from participants are kept in a secure place at the CeGIDD during the duration of the trial and will be destroyed at the end of the study. The final raw dataset will be accessible only to the sponsor (CHU) and the chief scientist's (SA) team. Anonymous data will be available to external parties upon approval of both the sponsor and the scientific committee. All publications will be made green or gold open access and the corresponding data will be provided as supplementary material or via a public repository (e.g. Zenodo), provided that there is no conflict with ethical guidelines.

Author contributions

Samuel Alizon, Carmen Lia Murall and Massical Rahmoun were the major contributors in the conception of the protocol. Samuel Alizon wrote the initial version of the manuscript. Christian

For peer review only

Selinger, Monique Baldellou, Claire Bernat, Marine Bonneau, Vanina Boué, Mathilde Buisson, Guillaume Christophe, Giuseppe D'Auria, Florence De Taroni, Vincent Foulongne, Rémy Froissart, Christelle Graf, Sophie Grasset, Soraya Groc, Christophe Hirtz, Audrey Jaussent, Frédérique Lorcy, Eric Picot, Marie-Christine Picot, Jacques Ravel, Jacques Reynes, Thérèse Rousset, Aziza Seddiki, Martine Teirlinck, Vincent Tribout, Édouard Tuaillon, Tim Waterboer, Nathalie Jacobs, Ignacio G Bravo, Michel Segondy and Natalie Boulle were involved in the conception of the protocol, in the implementation of the study and read and approved the final manuscript.

Ethics approval and consent to participate

The PAPCLEAR trial obtained favourable opinions from the Comité de Protection des Personnes (CPP) Sud Méditerranée I on May 11, 2016 (CPP number 16 42, reference number ID RCB 2016-A00712-49); from the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (CCTIRS) on July 12, 2016 (reference number 16.504); and from the Commission Nationale Informatique et Libertés (CNIL) on Dec 16, 2016 (reference number MMS/ABD/AR1612278, decision number DR-2016-488). This trial was authorised by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM) on July 20, 2016 (reference 20160072000007).

The protocol has been modified since its initial version and the latest modification was accepted by the CPP on Dec 12, 2018.

All participants in the study will sign an informed consent form prior to participation.

Acknowledgements

We thank all the study participants and the CeGIDD staff for their commitment to the study. We also thank reviewers and, in particular, Dr. Andrew Brouwer for his meticulous reading of the manuscript.

For peer review only

References

1. Tota, J.E., Chevarie-Davis, M., Richardson, L.A., Devries, M., Franco, E.L.: Epidemiology and burden of HPV infection and related diseases: implications for prevention strategies. Prev Med **53 Suppl 1**, 12–21 (2011). doi:10.1016/j.ypmed.2011.08.017

2. Monsonego, J., Zerat, L., Syrjänen, K., Zerat, J.C., Smith, J.S., Halfon, P.: Prevalence of genotype-specific HPV infection among women in France: implications for screening and vaccination. Gynecol Obstet Fertil **41**(5), 305–313 (2013). doi:10.1016/j.gyobfe.2013.03.003

3. Brun-Micaleff, E., Coffy, A., Rey, V., Didelot, M.-N., Combecal, J., Doutre, S., Daurès, J.-P., Segondy, M., Boulle, N.: Cervical cancer screening by cytology and human papillomavirus testing during pregnancy in french women with poor adhesion to regular cervical screening. J Med Virol **86**(3), 536–45 (2014). doi:10.1002/jmv.23764

4. Bruni, L., Diaz, M., Castellsagué, X., Ferrer, E., Bosch, F.X., de Sanjosé, S.: Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. J Infect Dis **202**(12), 1789–99 (2010). doi:10.1086/657321

5. Insinga, R.P., Dasbach, E.J., Elbasha, E.H., Liaw, K.-L., Barr, E.: Incidence and duration of cervical human papillomavirus 6, 11, 16, and 18 infections in young women: an evaluation from multiple analytic perspectives. Cancer Epidemiol Biomarkers Prev **16**(4), 709–15 (2007). doi:10.1158/1055-9965.EPI-06-0846

6. Woodman, C.B.J., Collins, S.I., Young, L.S.: The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer **7**(1), 11–22 (2007). doi:10.1038/nrc2050

7. Rodríguez, A.C., Schiffman, M., Herrero, R., Wacholder, S., Hildesheim, A., Castle, P.E., Solomon, D., Burk, R., Proyecto Epidemiológico Guanacaste Group: Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections. J Natl Cancer Inst **100**(7), 513–7 (2008). doi:10.1093/jnci/djn044

For peer review only

8. Trottier, H., Mahmud, S., Prado, J.C.M., Sobrinho, J.S., Costa, M.C., Rohan, T.E., Villa, L.L., Franco, E.L.: Type-Specific Duration of Human Papillomavirus Infection: Implications for Human Papillomavirus Screening and Vaccination. J Infect Dis **197**(10), 1436–1447 (2008). doi:10.1086/587698

9. Ramanakumar, A.V., Naud, P., Roteli-Martins, C.M., de Carvalho, N.S., de Borba, P.C., Teixeira, J.C., Blatter, M., Moscicki, A.-B., Harper, D.M., Romanowski, B., Tyring, S.K., Ramjattan, B., Schuind, A., Dubin, G., Franco, E.L.: Incidence and duration of type-specific human papillomavirus infection in high-risk HPV-naïve women: results from the control arm of a phase II HPV-16/18 vaccine trial. BMJ Open **6**(8), 011371 (2016). doi:10.1136/bmjopen-2016-011371

10. Houlihan, C.F., Baisley, K., Bravo, I.G., Kapiga, S., de Sanjosé, S., Changalucha, J., Ross, D.A., Hayes, R.J., Watson-Jones, D.: Rapid acquisition of HPV around the time of sexual debut in adolescent girls in Tanzania. Int J Epidemiol **45**(3), 762–773 (2016). doi:10.1093/ije/dyv367

11. Alizon, S., Murall, C.L., Bravo, I.G.: Why Human Papillomavirus Acute Infections Matter. Viruses **9**(10), 293 (2017). doi:10.3390/v9100293

12. Herrero, R., Wacholder, S., Rodríguez, A.C., Solomon, D., González, P., Kreimer, A.R., Porras, C., Schussler, J., Jiménez, S., Sherman, M.E., Quint, W., Schiller, J.T., Lowy, D.R., Schiffman, M., Hildesheim, A., Costa Rica Vaccine Trial Group: Prevention of persistent human papillomavirus infection by an HPV16/18 vaccine: a community-based randomized clinical trial in Guanacaste, Costa Rica. Cancer Discov **1**(5), 408–19 (2011). doi:10.1158/2159-8290.CD-11-0131

13. Canini, L., Perelson, A.S.: Viral kinetic modeling: state of the art. J Pharmacokinet Pharmacodyn **41**(5), 431–443 (2014). doi:10.1007/s10928-014-9363-3

14. Stanley, M.: Immune responses to human papillomavirus. Vaccine **24**(S1), 16–22 (2006).

For peer review only

doi:10.1016/j.vaccine.2005.09.002

15. Ferenczy, A., Franco, E.: Persistent human papillomavirus infection and cervical neoplasia. Lancet Oncol **3**(1), 11–6 (2002)

16. zur Hausen, H.: Review: Papillomaviruses — to Vaccination and Beyond. Biochemistry **73**(5), 498–503 (2008). doi:10.1134/S0006297908050027

17. Einstein, M.H., Schiller, J.T., Viscidi, R.P., Strickler, H.D., Coursaget, P., Tan, T., Halsey, N., Jenkins, D.: Clinician's guide to human papillomavirus immunology: knowns and unknowns. Lancet Infect Dis **9**(6), 347–56 (2009). doi:10.1016/S1473-3099(09)70108-2

18. Van Hede, D., Langers, I., Delvenne, P., Jacobs, N.: Origin and immunoescape of uterine cervical cancer. Presse Med **43**(12P2), 413–421 (2014). doi:10.1016/j.lpm.2014.09.005

19. Stanley, M.: Immunology of HPV infection. Curr Obstet Gynecol Rep **4**(4), 195–200 (2015). doi:10.1007/s13669-015-0134-y. Accessed 2017-03-20

20. Gao, W., Weng, J., Gao, Y., Chen, X.: Comparison of the vaginal microbiota diversity of women with and without human papillomavirus infection: a cross-sectional study. BMC Infect Dis **13**, 271 (2013). doi:10.1186/1471-2334-13-271

21. Brotman, R.M., Shardell, M.D., Gajer, P., Tracy, J.K., Zenilman, J.M., Ravel, J., Gravitt, P.E.: Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. J Infect Dis **210**(11), 1723–33 (2014). doi:10.1093/infdis/jiu330

22. Koutsky, L.A., Ault, K.A., Wheeler, C.M., Brown, D.R., Barr, E., Alvarez, F.B., Chiacchierini, L.M., Jansen, K.U., Proof of Principle Study Investigators: A controlled trial of a human papillomavirus type 16 vaccine. N Engl J Med **347**(21), 1645–51 (2002). doi:10.1056/NEJMoa020586

23. Riethmuller, D., Jacquard, A.-C., Lacau St Guily, J., Aubin, F., Carcopino, X., Pradat, P., Dahlab, A., Prétet, J.-L.: Potential impact of a nonavalent hpv vaccine on the occurrence of hpv-

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

related diseases in france. BMC Public Health **15**, 453 (2015). doi:10.1186/s12889-015-1779-1

24. Joura, E.A., Giuliano, A.R., Iversen, O.-E., Bouchard, C., Mao, C., Mehlsen, J., Moreira, E.D. Jr, Ngan, Y., Petersen, L.K., Lazcano-Ponce, E., Pitisuttithum, P., Restrepo, J.A., Stuart, G., Woelber, L., Yang, Y.C., Cuzick, J., Garland, S.M., Huh, W., Kjaer, S.K., Bautista, O.M., Chan, I.S.F., Chen, J., Gesser, R., Moeller, E., Ritter, M., Vuocolo, S., Luxembourg, A., Broad Spectrum HPV Vaccine Study: A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med **372**(8), 711–23 (2015). doi:10.1056/NEJMoa1405044

25. Murall, C.L., Bauch, C.T., Day, T.: Could the human papillomavirus vaccines drive virulence evolution? Proc Biol Sci **282**, 20141069 (2015). doi:10.1098/rspb.2014.1069

26. Alizon, S., Méthot, P.-O.: Reconciling Pasteur and Darwin to control infectious diseases. PLoS Biol **16**(1), 2003815 (2018). doi:10.1371/journal.pbio.2003815

27. Moscicki, A.-B., Ma, Y., Wibbelsman, C., Darragh, T.M., Powers, A., Farhat, S., Shiboski, S.: Rate of and Risks for Regression of CIN-2 in adolescents and young women. Obstet Gynecol **116**(6), 1373–1380 (2010). doi:10.1097/AOG.0b013e3181fe777f

28. Buck Jr., H.W.: Warts (genital). BMJ Clin Evid **2015**, 1602 (2015)

29. Herrero, R., González, P., Markowitz, L.E.: Present status of human papillomavirus vaccine development and implementation. Lancet Oncol **16**(5), 206–16 (2015). doi:10.1016/S1470-2045(14)70481-4

30. Maver, P.J., Poljak, M.: Progress in prophylactic human papillomavirus (HPV) vaccination in 2016: A literature review. Vaccine (2018). doi:10.1016/j.vaccine.2017.07.113

31. Fagot, J.-P., Boutrelle, A., Ricordeau, P., Weill, A., Allemand, H.: HPV vaccination in France: uptake, costs and issues for the National Health Insurance. Vaccine **29**(19), 3610–6 (2011). doi:10.1016/j.vaccine.2011.02.064

For peer review only

32. Ben Hadj Yahia, M.-B., Dervaux, B., Duport, N., Floret, D., Gaillot, J., Heard, I., Jacquet, A., Le Goaster, C., Lévy-Bruhl, D., Morer, I., Parent du Chatelet, I., Peigue-Lafeuille, H., Rumeau-Pichon, C.: Vaccination contre les infections à papilloamvirus. Technical report, Haut Conseil de la Santé Publique, Paris, France (2014). https://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=454

33. Clarivet, B., Picot, E., Marchandin, H., Tribout, V., Rachedi, N., Schwartzentruber, E., Ledésert, B., Dereure, O., Guillot, B., Picot, M.-C.: Prevalence of Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium in asymptomatic patients under 30 years of age screened in a French sexually transmitted infections clinic. Eur J Dermatol **24**(5), 611–6 (2014). doi:10.1684/ejd.2014.2413

34. Winer, R.L., Hughes, J.P., Feng, Q., O'Reilly, S., Kiviat, N.B., Holmes, K.K., Koutsky, L.A.: Condom use and the risk of genital human papillomavirus infection in young women. N Engl J Med **354**(25), 2645–54 (2006). doi:10.1056/NEJMoa053284

35. Winer, R.L., Hughes, J.P., Feng, Q., Stern, J.E., Xi, L.F., Koutsky, L.A.: Incident Detection of High-Risk Human Papillomavirus Infections in a Cohort of High-Risk Women Aged 25-65 Years. J Infect Dis **214**(5), 665–75 (2016). doi:10.1093/infdis/jiw074

36. Ravel, J., Brotman, R.M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D.W., Sakamoto, J., Koenig, S.S., Fu, L., Zhou, X., Hickey, R.J., Schwebke, J.R., Forney, L.J.: Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. Microbiome **1**(1), 29 (2013). doi:10.1186/2049-2618-1-29

37. Kleter, B., van Doorn, L.-J., ter Schegget, J., Schrauwen, L., van Krimpen, K., Burger, M.,

For peer review only

ter Harmsel, B., Quint, W.: Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses. Am J Pathol **153**(6), 1731–1739 (1998). doi:10.1016/S0002-9440(10)65688-X

38. Geraets, D.T., Struijk, L., Kleter, B., Molijn, A., van Doorn, L.-J., Quint, W.G.V., Colau, B.: The original SPF10 LiPA25 algorithm is more sensitive and suitable for epidemiologic HPV research than the SPF10 INNO-LiPA Extra. J Virol Meth **215-216**, 22–29 (2015). doi:10.1016/j.jviromet.2015.01.001

39. Gravitt, P.E., Peyton, C.L., Alessi, T.Q., Wheeler, C.M., Coutlée, F., Hildesheim, A., Schiffman, M.H., Scott, D.R., Apple, R.J.: Improved amplification of genital Human Papillomaviruses. J Clin Microbiol **38**(1), 357–361 (2000)

40. Micalessi, I.M., Boulet, G.A.V., Bogers, J.J., Benoy, I.H., Depuydt, C.E.: High-throughput detection, genotyping and quantification of the human papillomavirus using real-time PCR. Clin Chem Lab Med **50**(4), 655–61 (2012). doi:10.1515/cclm.2011.835

41. Hunter, P.J., Sheikh, S., David, A.L., Peebles, D.M., Klein, N.: Cervical leukocytes and spontaneous preterm birth. Journal of Reproductive Immunology **113**, 42–49 (2016). doi:10.1016/j.jri.2015.11.002

42. Shannon, B., Yi, T.J., Perusini, S., Gajer, P., Ma, B., Humphrys, M.S., Thomas-Pavanel, J., Chieza, L., Janakiram, P., Saunders, M., Tharao, W., Huibner, S., Shahabi, K., Ravel, J., Rebbapragada, A., Kaul, R.: Association of HPV infection and clearance with cervicovaginal immunology and the vaginal microbiota. Mucosal Immunology **10**(5), 1310–1319 (2017). doi:10.1038/mi.2016.129

43. Lajoie, J., Juno, J., Burgener, A., Rahman, S., Mogk, K., Wachihi, C., Mwanjewe, J., Plummer, F.A., Kimani, J., Ball, T.B., Fowke, K.R.: A distinct cytokine and chemokine profile at the genital mucosa is associated with HIV-1 protection among HIV-exposed seronegative commercial sex workers. Mucosal Immunol **5**(3), 277–287 (2012). doi:10.1038/mi.2012.7

For peer review only

44. Juno, J.A., Boily-Larouche, G., Lajoie, J., Fowke, K.R.: Collection, isolation, and flow cytometric analysis of human endocervical samples. J Vis Exp **89**, 51906 (2014). doi:10.3791/51906

45. Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J.: Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol **74**(8), 2461 (2008). doi:10.1128/AEM.02272-07

46. Findley, K., Oh, J., Yang, J., Conlan, S., Deming, C., Meyer, J.A., Schoenfeld, D., Nomicos, E., Park, M., NIH Intramural Sequencing Center Comparative Sequencing Program, Kong, H.H., Segre, J.A.: Topographic diversity of fungal and bacterial communities in human skin. Nature **498**(7454), 367–370 (2013). doi:10.1038/nature12171. Accessed 2017-09-13

47. Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., Forney, L.J.: Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA **108**, 4680–7 (2011). doi:10.1073/pnas.1002611107

48. Gajer, P., Brotman, R.M., Bai, G., Sakamoto, J., Schütte, U.M.E., Zhong, X., Koenig, S.S.K., Fu, L., Ma, Z.S., Zhou, X., Abdo, Z., Forney, L.J., Ravel, J.: Temporal dynamics of the human vaginal microbiota. Sci Transl Med **4**(132), 132–52 (2012). doi:10.1126/scitranslmed.3003605

49. Johne, R., Müller, H., Rector, A., van Ranst, M., Stevens, H.: Rolling-circle amplification of viral DNA genomes using phi29 polymerase. Trends Microbiol **17**(5), 205–211 (2009). doi:10.1016/j.tim.2009.02.004

50. Nowak, M.A., May, R.M.: Virus Dynamics: Mathematical Principles of Immunology and Virology. Oxford University Press, Oxford, USA (2000)

51. Stafford, M.A., Corey, L., Cao, Y., Daar, E.S., Ho, D.D., Perelson, A.S.: Modeling plasma virus concentration during primary HIV infection. J. theor. Biol. **203**(3), 285–301 (2000).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

doi:10.1006/jtbi.2000.1076

52. Perelson, A.S.: Modelling viral and immune system dynamics. Nat. Rev. Immunol. **2**(1), 28–36 (2002). doi:10.1038/nri700

53. Murall, C.L., Jackson, R., Zehbe, I., Boulle, N., Segondy, M., Alizon, S.: Epithelial stratification shapes infection dynamics. PLoS Comput Biol **15**(1), 1006646 (2019). doi:10.1371/journal.pcbi.1006646

54. Steimer, J.L., Vozeh, S., Racine Poon, A., Holford, N., O'Neil, R.: The population approach: rationale, methods and applications in clinical pharmacology and drug development. In: Balant, P.G.W..L. (ed.) Handbook of Experimental Pharmacology, vol. 110, pp. 405–451. Springer, Berlin (1994)

55. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. Journal of Statistical Software **67**(1) (2015). doi:10.18637/jss.v067.i01

56. Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., Deng, L., Yeliseyev, V., Delaney, M.L., Liu, Q., Olle, B., Stein, R.R., Honda, K., Bry, L., Gerber, G.K.: MDSINE: Microbial dynamical systems INference engine for microbiome time-series analyses. Genome Biology **17**, 121 (2016). doi:10.1186/s13059-016-0980-6. Accessed 2017-03-09

57. Faust, K., Lahti, L., Gonze, D., de Vos, W.M., Raes, J.: Metagenomics meets time series analysis: unraveling microbial community dynamics. Curr Opin Microbiol **25**, 56–66 (2015). doi:10.1016/j.mib.2015.04.004

58. Fox, G.A., Negrete-Yankelevich, S., Sosa, V.J.: Ecological Statistics: Contemporary Theory and Application. Oxford University Press, Oxford, USA (2015)

59. Shi, Y., Li, L., Hu, Z., Li, S., Wang, S., Liu, J., Wu, C., He, L., Zhou, J., Li, Z., Hu, T., Chen, Y., Jia, Y., Wang, S., Wu, L., Cheng, X., Yang, Z., Yang, R., Li, X., Huang, K., Zhang, Q., Zhou, H., Tang, F., Chen, Z., Shen, J., Jiang, J., Ding, H., Xing, H., Zhang, S., Qu, P., Song, X., Lin,

For peer review only

Z., Deng, D., Xi, L., Lv, W., Han, X., Tao, G., Yan, L., Han, Z., Li, Z., Miao, X., Pan, S., Shen, Y., Wang, H., Liu, D., Gong, E., Li, Z., Zhou, L., Luan, X., Wang, C., Song, Q., Wu, S., Xu, H., Shen, J., Qiang, F., Ma, G., Liu, L., Chen, X., Liu, J., Wu, J., Shen, Y., Wen, Y., Chu, M., Yu, J., Hu, X., Fan, Y., He, H., Jiang, Y., Lei, Z., Liu, C., Chen, J., Zhang, Y., Yi, C., Chen, S., Li, W., Wang, D., Wang, Z., Di, W., Shen, K., Lin, D., Shen, H., Feng, Y., Xie, X., Ma, D.: A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. Nat Genet **45**(8), 918–22 (2013). doi:10.1038/ng.2687

60. Chen, D., Gaborieau, V., Zhao, Y., Chabrier, A., Wang, H., Waterboer, T., Zaridze, D., Lissowska, J., Rudnai, P., Fabianova, E., Bencko, V., Janout, V., Foretova, L., Mates, I.N., Szeszenia-Dabrowska, N., Boffetta, P., Pawlita, M., Lathrop, M., Gyllensten, U., Brennan, P., McKay, J.D.: A systematic investigation of the contribution of genetic variation within the MHC region to HPV seropositivity. Hum Mol Genet **24**(9), 2681–2688 (2015). doi:10.1093/hmg/ddv015

61. Chen, D., Gyllensten, U.: Lessons and implications from association studies and post-GWAS analyses of cervical cancer. Trends Genet **31**(1), 41–54 (2015). doi:10.1016/j.tig.2014.10.005

62. Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., Cozzi-Lepri, A., De Luca, A., Easterbrook, P., Francioli, P., Mallal, S., Martinez-Picado, J., Miro, J.M., Obel, N., Smith, J.P., Wyniger, J., Descombes, P., Antonarakis, S.E., Letvin, N.L., McMichael, A.J., Haynes, B.F., Telenti, A., Goldstein, D.B.: A whole-genome association study of major determinants for host control of HIV-1. Science **317**(5840), 944–947 (2007). doi:10.1126/science.1143767

63. Huber, P.J.: The 1972 Wald Lecture Robust Statistics: A Review. Ann Math Stat **43**(4), 1041–1067 (1972). doi:10.1214/aoms/1177692459

64. Winer, R.L., Feng, Q., Hughes, J.P., O'Reilly, S., Kiviat, N.B., Koutsky, L.A.: Risk of female human papillomavirus acquisition associated with first male sex partner. J Infect Dis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

197(2), 279–82 (2008). doi:10.1086/524875

65. Herrero, R.: Human Papillomavirus (HPV) Vaccines: Limited Cross-Protection against Additional HPV Types. J Infect Dis **199**(7), 919–922 (2009). doi:10.1086/597308

66. Depuydt, C.E., Verstraete, L., Berth, M., Beert, J., Bogers, J.-P., Salembier, G., Vereecken, A.J., Bosmans, E.: Human papillomavirus positivity in women undergoing intrauterine insemination has a negative effect on pregnancy rates. Gynecol Obstet Invest **81**(1), 41–6 (2016). doi:10.1159/000434749

67. WHO HPV LabNet. World Health Organization. http://www.who.int/biologicals/areas/human_papillomavirus/WHO_HPV_LabNet/en/

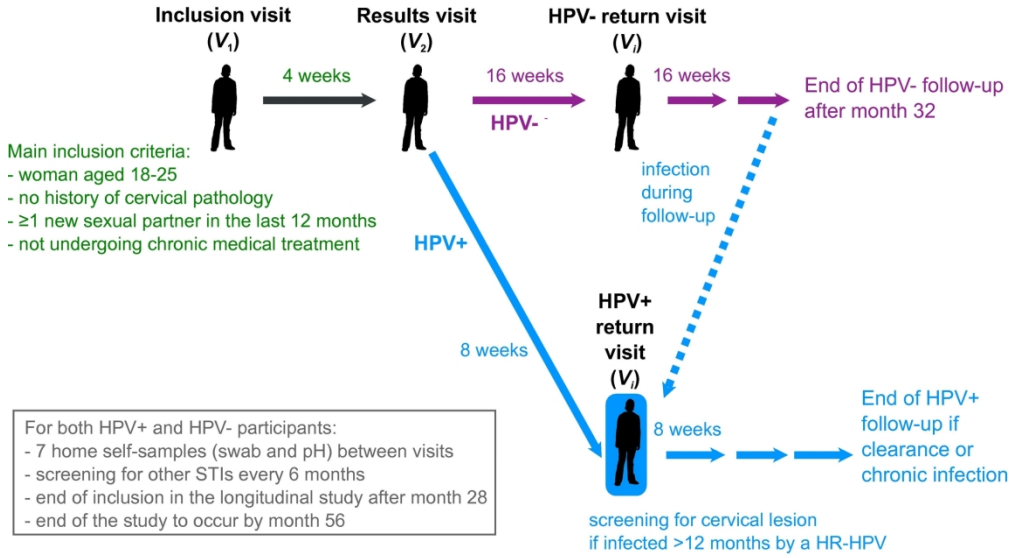
For peer review only

Figure captions

Figure 1: General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V_1), a results visit (V_2) and then return visits (V_i with $i > 2$). For the cross-sectional study, participants only have V_1 and V_2.

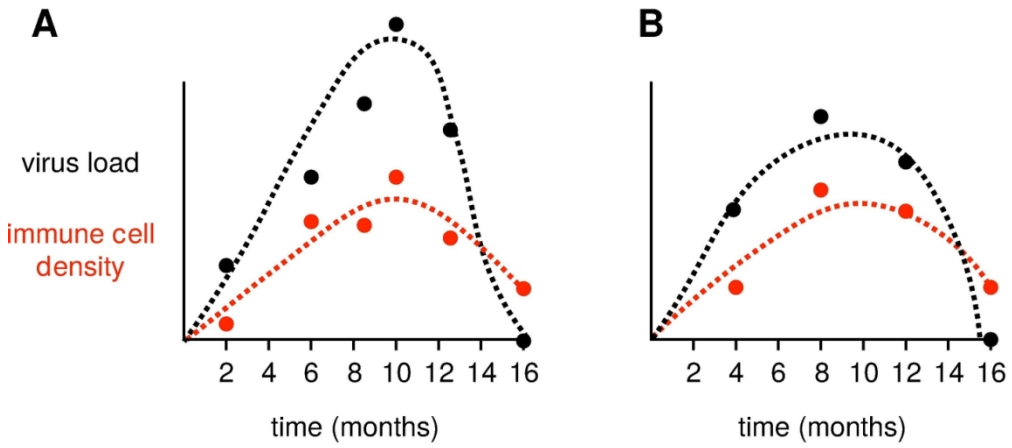
Figure 2: Fitting viral kinetics models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

For peer review only



General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V1), a results visit (V2) and then return visits (Vi with i > 2). For the cross-sectional study, participants only have V1 and V2.

190x104mm (300 x 300 DPI)



Fitting kinetics dynamical models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

120x52mm (300 x 300 DPI)

BMJ Open

The natural history, dynamics, and ecology of Human papillomaviruses in genital infections of young women: protocol of the PAPCLEAR cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025129.R2
Article Type:	Protocol
Date Submitted by the Author:	09-Apr-2019
Complete List of Authors:	Murall, Carmen Lia; Centre National de la Recherche Scientifique, MIVEGEC Rahmoun, Massilva; Centre National de la Recherche Scientifique, MIVEGEC Selinger, Christian; Centre National de la Recherche Scientifique, MIVEGEC Baldellou, Monique; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Bernat, Claire; Centre National de la Recherche Scientifique, MIVEGEC Bonneau, Marine; Centre Hospitalier Regional Universitaire de Montpellier, Department of Obstetrics and Gynaecology Boué, Vanina; Centre National de la Recherche Scientifique, MIVEGEC Buisson, Mathilde; Centre Hospitalier Regional Universitaire de Montpellier, Department of Research and Innovation Christophe, Guillaume; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD D'Auria, Giuseppe; Fundacio per al Foment de la Investigacio Sanitaria i Biomedica, Sequencing and Bioinformatics Service; Centro de Investigacion Biomedica en Red de Epidemiologia y Salud Publica De Taroni, Florence; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Foulongne, Vincent; Centre Hospitalier Regional Universitaire de Montpellier, Department of Bacteriology and Virology Froissart, Rémy; Centre National de la Recherche Scientifique, MIVEGEC Graf, Christelle; Centre Hospitalier Regional Universitaire de Montpellier, Department of Obstetrics and Gynaecology Grasset, Sophie; Centre Hospitalier Regional Universitaire de Montpellier, Department of Virology; Centre National de la Recherche Scientifique, MIVEGEC Groc, Soraya; Centre Hospitalier Regional Universitaire de Montpellier, Department of Virology; Centre National de la Recherche Scientifique, MIVEGEC Hirtz, Christophe; Centre Hospitalier Regional Universitaire de Montpellier, LBPC/PPC, IRMB Jaussent, Audrey; Centre Hospitalier Regional Universitaire de Montpellier, Department of Medical Information Lajoie, Julie; University of Manitoba College of Medicine, Department of Medical microbiology Lorcy, Frédérique; Centre Hospitalier Regional Universitaire de

	Montpellier, Laboratoire d'anatomie et cytologie pathologiques Picot, Eric; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD PICOT, Marie-Christine; Centre Hospitalier Regional Universitaire de Montpellier, Department of Medical Information Ravel, Jacques; University of Maryland School of Medicine, Institute for Genome Sciences Reynes, Jacques; Centre Hospitalier Regional Universitaire de Montpellier, Department of Infectious and Tropical Diseases Rousset, Thérèse; Centre Hospitalier Regional Universitaire de Montpellier, Department of pathology and oncobiology Seddiki, Aziza; Centre Hospitalier Regional Universitaire de Montpellier, Department of Research and Innovation Teirlinck, Martine; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Tribout, Vincent; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Tuaillon, Edouard; Centre Hospitalier Regional Universitaire de Montpellier, Department of bacteriology and virology Waterboer, Tim; Deutsches Krebsforschungszentrum, Infections and Cancer Epidemiology Jacobs, Nathalie; Universite de Liege Faculte des Sciences, GIGA-Research, Cellular and molecular immunology Bravo, Ignacio; Centre National de la Recherche Scientifique, MIVEGEC Segondy, Michel; Centre Hospitalier Regional Universitaire de Montpellier, Department of Bacteriology and Virology Boulle, Nathalie; Centre Hospitalier Regional Universitaire de Montpellier, Department of pathology and oncobiology Alizon, Samuel; Centre National de la Recherche Scientifique, MIVEGEC
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology, Immunology (including allergy), Genetics and genomics
Keywords:	VIROLOGY, IMMUNOLOGY, Epidemiology < INFECTIOUS DISEASES, MICROBIOLOGY, GENETICS

SCHOLARONE™
Manuscripts

STUDY PROTOCOL

The natural history, dynamics, and ecology of Human papillomaviruses in genital infections of young women: protocol of the PAPCLEAR cohort study

Carmen Lía Murall[1] , Massilva Rahmoun[1], Christian Selinger[1], Monique Baldellou[2], Claire Bernat[1], Marine Bonneau[3] , Vanina Boué[1] , Mathilde Buisson[4], Guillaume Christophe[2], Giuseppe D'Auria[5,6], Florence De Taroni[2], Vincent Foulongne[7,8], Rémy Froissart[1], Christelle Graf[3], Sophie Grasset[1,2], Soraya Groc[1,7], Christophe Hirtz[9], Audrey Jaussent[10], Julie Lajoie[11], Frédérique Lorcy[12], Eric Picot[2], Marie-Christine Picot[10], Jacques Ravel[13], Jacques Reynes[11], Thérèse Rousset[12], Aziza Seddiki[4], Martine Teirlinck[2], Vincent Tribout[2], Édouard Tuaillon[8], Tim Waterboer[15] , Nathalie Jacobs[16], Ignacio G Bravo[1], Michel Segondy[7,8], Nathalie Boulle[8,12] and Samuel Alizon[1,*]

For peer review only

Word count: 5265 words, excluding title page, abstract, references, figures and tables.

1 Laboratoire MIVEGEC (UMR 5290 CNRS, IRD, UM), 911, avenue Agropolis, BP 64501, 34394 Montpellier, France
2 Center for Free Information, Screening and Diagnosis (CeGIDD), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
3 Department of Obstetrics and Gynaecology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France.
4 Department of Research and Innovation (DRI), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
5 Sequencing and Bioinformatics Service, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública), 46020 Valencia, Spain
6 CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain
7 Department of Virology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
8 Pathogenesis and Control of Chronic Infections, INSERM, CHU, University of Montpellier, Montpellier, France
9 University of Montpellier, LBPC/PPC- IRMB, CHU de Montpellier, 80 rue Augustin Fliche, Montpellier, France
10 Department of Medical Information (DIM), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
11 Department of Medical microbiology, University of Manitoba, 745 Bannatyne, Winnipeg, Canada
12 Department of pathology and oncobiology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
13 Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, USA
14 Department of Infectious and Tropical Diseases, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
15 German Cancer Research Center (DKFZ), Infections and Cancer Epidemiology, Im Neuenheimer Feld 280, Heidelberg, Germany
16 GIGA-Research, Cellular and molecular immunology, University of Liège, 3 Avenue de l'Hôpital, 4000 Liège, Belgium

* Author for correspondence: samuel.alizon@cnrs.fr

For peer review only

Abstract

Introduction

Human papillomaviruses (HPVs) are responsible for one third of all cancers caused by infections. Most HPV studies focus on chronic infections and cancers, and thus, we know little about the early stages of viral infection. In particular, the effects of the dynamic interactions between the immune system, the microbiota, and the viral and host genetics on infection clearance or persistence remains poorly understood.

Methods and Analysis

We follow 150 women, aged 18-25 years, longitudinally to monitor immune response features (cytokines and immune cells in the genital tract, circulating anti-HPV antibodies), virus load of HPVs, and vaginal microbiota composition. This is complemented by the assessment of viral and human genetics and behavioural data. To increase the statistical power of the epidemiological arm of the study, an additional 150 women are screened cross-sectionally.

Ethics and Dissemination

This study has been approved by the Comité de Protection des Personnes Sud Méditerranée I (reference number 2016-A00712-49); by the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (reference number 16.504); by the Commission Nationale Informatique et Libertés (reference number MMS/ABD/AR1612278, decision number DR-2016-488) and by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (reference 20160072000007). The results will be published in preprint servers, post-print servers, peer-reviewed journals and disseminated through conferences.

Trial registration number: NCT02946346

Keywords: HPV; acute infection; persistence; virus load; immunity; microbiota; viral kinetics

For peer review only

Article summary

Strengths and limitations of this study

- A dense follow-up with visits every two months for infected women and additional self-sampling every week.

- The combination of virological (virus load), immunological (cytokine concentrations and immune cell percentages) and environmental (vaginal microbiota composition, pH) measurements at each visit.

- A limitation is that the density of the follow-up limits the number of participants (N=150), which can affect the power of epidemiological analyses.

- We complement the longitudinal study with a cross-sectional study of N=150 women to increase statistical power.

For peer review only

Introduction

Epidemiology of HPV genital infections in young adults and public health implications

Infections by Human Papillomaviruses (HPVs) are likely the most common sexually transmitted infection (STI) globally. It is often estimated that, worldwide, more than 80% of sexually-active individuals will be infected by an HPV type [1]. In France, a study performed in 2013 in the Paris area estimated the prevalence of HPV genital infections to be 25% in women below 25 years of age [2]. In the area of Montpellier (France), the prevalence of oncogenic HPVs (also referred to as 'high-risk', HR, HPVs) in pregnant women aged 16 to 42 years was close to 20% [3]. These numbers are consistent with worldwide estimates according to which HPVs are most prevalent in women under 25 years of age, with an estimated overall prevalence of 24% [4].

Fortunately, the vast majority of infections by HPVs are asymptomatic and benign. Even for HPV16, which is probably the most oncogenic human virus, only a minority of infections (less than 10%) become persistent [5], and then a minority of these (12%) progress to cancer if untreated [1, 6]. Indeed, it is estimated that approximately 70 to 100% of infections by HPVs are cleared within 12 to 24 months, with strong differences between virus types [5, 7–9]. Recent studies suggest that primo-infections could be shorter in young girls [10] but, in general, there are many unknowns about the biology of non-persisting infections [11].

Our lack of knowledge partly comes from the fact that in vaccine trials, from which most of the data on infection duration originate, participants are followed every six months for several years [5, 7, 9, 12]. This frequency is sufficient to estimate the time to clearance (or persistence) but it is not precise enough to understand the within-host dynamics, often referred to as 'kinetics' [13], of infections that last on average 6 to 24 months. Arbitrarily, after 24 months, an infection is often considered to be persistent [14].

Some factors have been shown to correlate with persistence (e.g. immunosuppression, smoking, and co-infection with other STIs [15]) but we do not know how these affect viral kinetics. Also, some changes in viral-immunity interactions appear to be related to persistence

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

and disease progression [16–19] but, again, we do not know the underlying interactions between the viruses, the host target cells, and the immune response in acute infections [11]. Finally, it has been argued that the vaginal microbiota may differ between HPV-infected and HPV-uninfected women [20] and that specific microbiota composition may interact with HPV detection [21]. However, it is difficult to disentangle the cause and the consequence. For instance, does the microbiota composition change after the establishment of an HPV infection, or do certain microbiota compositions increase susceptibility to HPV infection?

A better understanding of the within-host infection dynamics and of the determinants of clearance and persistence of viral infection is particularly important in the context of vaccination [22–25]. Indeed, the long-term efficacy of the anti-HPVs vaccines at the population level will largely depend on the within-host viral dynamics because, ultimately, most selective pressures on viral populations occur via the immune response [26]. Furthermore, a better understanding of acute HPV infections can shed new light on issues related to latency, fertility, or immunotherapies [11].

Prevention strategies and treatment

Treatment

Since most infections by HPVs are benign in young adults and clear within six to 24 months, the current standard of care is to avoid over-treatment, even in the presence of cervical lesions [27]. Clinical interventions (colposcopies, biopsies, and surgery) are less often performed with young women (< 25 years) and only for high-grade (pre-cancerous) lesions (cervical intraepithelial neoplasia grade 2, CIN-2, or more). Low-grade lesions (CIN-1) are not systematically treated but rather monitored yearly to detect any progression to high-grade lesions.

Genital warts caused by non-oncogenic HPVs (often referred to as ‘low-risk’, LR, HPVs) can be removed by surgery or treated with bi- and trichloroacetic acid, cryotherapy or other treatments [28].

For peer review only

HPV vaccination

There are currently three licensed vaccines: a bivalent vaccine (Cervarix®) targeting HPV16 and HPV18 (together accounting for 70% of cervical cancers [1]), a quadrivalent vaccine (Gardasil®) that additionally targets HPV6 and HPV11 (non-oncogenic, but highly prevalent and associated to benign proliferative lesions) and, since 2014, a nonavalent vaccine (Gardasil 9®) that targets five more oncogenic types (HPV31, HPV33, HPV45, HPV52, and HPV58, which altogether account for 20% of cervical cancers [24]). These vaccines succeed in eliciting a protective immune response against new infections by the targeted viruses, and are used throughout the world, albeit with wide variation in coverage (for reviews, see e.g. [29, 30]).

Vaccination campaigns in France started in 2006 but with limited coverage: it reached 28.5% in 2008 [31] and has been decreasing ever since [32]. The vaccine is recommended for girls from 11 to 14 years of age, currently with a vaccination scheme of two doses with a six months interval. A catch-up is organised for girls aged 15-19 years, with a three-doses vaccination scheme. Vaccination is reimbursed by the French Social Security but is not mandatory. It is also recommended for men who have sex with men (MSM) as well as for immuno-compromised people [32]. Vaccination is now the primary prevention strategy against cervical cancers.

Screening

In France, the secondary prevention strategy against cervical cancer is routine individual cytology-based screening for pre-cancerous and cancerous cervical lesions in women between 25 and 65 years old. Cytology can also be performed in younger women if they report risk factors for cervical cancer (multiple partners, chronic STIs or HIV infection [32]). Detection of oncogenic HPVs is proposed for triage in case of abnormal cytology (i.e. high-grade or low-grade squamous intraepithelial lesion, HSIL and LSIL respectively, or Atypical Squamous Cells of Undetermined Significance, ASCUS).

Primary objectives

The first primary objective of this cohort study is to decipher the kinetics and ecology of cervical HPV infections in healthy young women, i.e. follow the population dynamics of the virus, the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

target epithelial cells, and the immune effectors.

The second primary objective is to characterise the diversity of genital HPVs in young women in the region of Montpellier in relationship with their lifestyle, vaccination status, vaginal microbiota, and human genetics.

Secondary objectives

A secondary objective is to characterise the acquisition and clearance dynamics of cervical HPV infections as a function of viral diversity, host immunity, vaginal microbiota and human genetics.

A final objective is to investigate variations in genetic diversity of HPVs during cervical infections.

Methods and analysis

Participants

The study population is composed of young women at risk of HPV infection. The age class was chosen because it exhibits high HPV prevalence (24% worldwide [4] and approximately 25% in France [2]). Inclusion of younger women would have raised technical issues because of the requirement for parental consent.

Women are recruited through a social media page, and through posters and leaflets distributed at the Universities in Montpellier and at the Montpellier STI screening centre (*Centre Gratuit d'Information de Dépistage et de Diagnostic*, CeGIDD). The composition of the population visiting the CeGIDD has already been documented in an earlier study [33]. In total, the centre is visited by approximately 3,000 women per year, the majority of which are under 25 years of age (80%). Approximately 40% of the attendants report three or more partners over the last twelve months and approximately 50% report using adequate behaviour for prevention against HIV.

For peer review only

Inclusion criteria

Participants are women from 18 to 25 years old living in the metropolitan area of Montpellier. They must be sexually active with at least one new partner over the last 12 months. This criteria is fixed to maximise the incidence of new HPV infections. Participants must be able to and willing to give written informed consent: they must sign an informed consent form, understand the requirements for the study, and be affiliated to a French social security scheme (which is a state requirement).

Women cannot be included in the study if they have a history of HPV-associated pathology (genital warts or cervical lesions), if they are pregnant or intending to become pregnant in the coming year, infected by HIV, undergoing (or planning to undergo) intense medical treatment (biotherapy, chemotherapy, immunosuppression), planning on moving outside the Montpellier metropolitan area within the next 18 months, in a dependency or employment with the sponsor or the investigator, if they participated in a clinical trial involving administration of drugs within the last four weeks or if they belong to a vulnerable group (e.g. children, adults with physical or mental disabilities).

Design/setting

This study has a longitudinal component aimed at deciphering within-host dynamics and a cross-sectional component aimed at understanding the diversity of HPV infections in young adults in the area of Montpellier, France. The general structure of the study is shown in Fig 1.

If a woman fits the main inclusion criteria, she can go through an inclusion visit ($V1$) with a physician (gynaecologist or midwife) at the CeGIDD. During this visit, the study investigator presents the study and checks all inclusion criteria before asking the participant to read and sign the informed consent form. Participants then undergo a medical consultation during which a number of samples are collected (see below). They then fill out health and lifestyle questionnaires and are given cotton-flocked swabs for self-sampling at home the next visit,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

along with instructions on how to fill in weekly questionnaires through an online form (these are performed throughout the study).

An appointment is scheduled four weeks later for the Results visit (V_2), where the cervical cytology results are communicated. Additional samples are collected and self-sample swabs for home collection are provided.

The next return visits (V_i, where $i > 2$) are as follows:

- Participants with a positive DEIA HPV test (see below), i.e. infected by an *Alphapapillomavirus*, at V_1 join the HPV positive (HPV+) arm of the study with return visits scheduled every 2 months.
- Participants with a negative DEIA HPV test at V_1 join the HPV negative (HPV-) arm with return visits scheduled every 4 months.
- HPV- participants infected by an *Alphapapillomavirus* move to the HPV+ arm.

Intervals between visits are based on earlier results showing that HPV infections last from 9 to 18 months on average depending on the HPV type [5, 7–9] and that a total follow-up of 4 months yields results that are difficult to analyse [21]. The longer interval in the HPV- arm is based on the estimated incidence for HPV genital infections in young women, which is greater than 30% [34, 35].

Participants in the HPV- arm are followed until month 32 of the study.

Participants in the HPV+ arm are followed until they clear the infection or until they have been infected for 24 months (after which we consider that the infection is persistent). Clearance is defined as being negative at two visits in a row for the first HPV type detected in the follow-up.

In between these visits to the CeGIDD, participants are asked to perform regular (every week for HPV+ and every second week for HPV-) self-samples using vaginal swabs, to measure vaginal pH and to fill in a short questionnaire. Self-samples are stored in the participants' freezer and brought back at every visit.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The study will end with the last HPV+ participant having cleared the infection or been infected for 24 months.

Patients and public involvement

Since all participants are healthy, they are referred to as participants rather than patients. As in any longitudinal study, ensuring participant commitment is challenging. To achieve this goal, we have set up a compensation of 40 EUR per visit and an additional 10 EUR in case of a complete follow-up. Furthermore, participants who have answered a sufficient number of questionnaires and brought back a sufficient number of self-samples get a 100 EUR bonus at the end. Overall, a participant performing 12 return visits would gain a total compensation of 650 EUR.

Participants did not play a role in the design of this study.

Results of the study will be disseminated to participants who have left the study and to the general public via an email newsletter in French.

Visits

The summary of the visit schedule and of the samples collected at each visit is shown in Table 1.

Inclusion visit (V_1)

This visit takes place at the CeGIDD and is scheduled by the Clinical Research Technician (CRT) via phone or email.

Women meet a study investigator, who explains the goals and requirements of the study and checks that the inclusion criteria are met. If so, after a general discussion, the informed consent forms are signed.

For peer review only

A female physician/midwife performs a general exam and then a gynaecological exam during which the following samples are taken:

- vaginal pH cotton swab (EcoCare™),
- vaginal swab (Copan ESwab™) in 1mL Amies liquid for DNA extraction and microbiota analysis,
- vaginal swab (Copan ESwab™) in 1mL of RNA preservation medium,
- ophthalmic sponge (Weck-cel®) to collect cervical secretions for cytokines analysis,
- cervical smear in 20mL of Thinprep® (Preservcyt® liquid) for HPV and HSV assays, and cytology evaluation.

Following the gynaecological consultation, the participant meets with a nurse to measure body temperature, blood pressure and draw 20mL of blood (a 5mL tube for SNPs sequencing, a 10mL tube for immunophenotyping and a 5mL tube for HPV antibody titration). For the longitudinal study, the nurse provides the participant with 3 self-sampling kits, 3 pH strips, a freezer box to bring back to the next visit, as well as instructions on how to perform the home sampling and store the samples in her personal freezer until the next visit.

If the participant has not been tested for a STI in the last 3 months, the nurse draws an additional blood tube of 5mL to test for STIs (HIV, HCV, HBV) and collects vaginal self-samples for chlamydiae and gonorrhea detection. Syphilis testing is prescribed to participants who meet the STI clinic's guidelines.

Finally, the participant meets with the CRT to fill in questionnaires #1 (inclusion visit) and #3 (home). The CRT answers any remaining questions, explains how to fill the home questionnaires (#3) and sets an appointment for the Results visit.

Results visit (*V*2)

During this visit, the participants are given the result of cervical lesion screening using the liquid cytology (normal, ASCUS, LSIL or HSIL). Participants with a high-grade lesion (HSIL) exit the study and are referred to the gynaecology service of the CHU of Montpellier.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

During this visit, the physician/midwife collects additional samples: 2 vaginal swabs for DNA and RNA analysis, and a cervical smear in 10mL of PBS (to confirm HPV status and perform flow cytometry analyses).

The participant fills in questionnaires #2 (for return visits) and #3 (home). An appointment for the next visit is set and swabs for home self-sampling are given.

Return visits (*Vi*)

These visits only occur in the longitudinal study.

HPV- arm. Participants uninfected by HPV visit the clinic every 4 months until month 26. During these visits, the same samples as in the inclusion visit (*V*1**)** are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

The nurse only draws blood if a screening test for STIs other than HPV is required. The participant then fills in questionnaires #2 and #3 and an appointment is set for the next visit in 16 weeks.

If an HPV infection is detected in the cervical smear collected during this visit, the participant moves to the HPV+ arm and the CRT contacts the participant to move her appointment forward.

HPV+ arm. Participants infected by HPV visit the clinic every 2 months. They cannot switch arm and will remain in the HPV+ arm until clearance or the end of the study. During the visits, the same samples as in the inclusion visit (*V*0) are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

The nurse then draws 5mL of blood for HPV antibody titration. If this is the first HPV+ visit following an HPV- visit, the nurse also draws 10mL of blood for immunophenotyping. Finally, if a test for additional STIs is needed, the nurse draws 5mL of blood and collects vaginal self-samples for STI detection.

Importantly, if the participant has been infected by a HR-HPV for more than 12 months and

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

cytology has not been performed within the last 12 months, the cervical smear is put in Thinprep® fixation medium, instead of PBS, for cytological analysis (cervical lesion screening).

Finally, the participant fills in questionnaires #2 and #3, receives self-samples for home collection and an appointment is set for the next visit in 8 weeks.

Endpoints

The primary endpoint for the study is the inclusion and follow-up of HPV-infected women in order to describe the kinetics of HPV virus load, and the associated immune response.

Secondary endpoints are the characterisation of the interactions between the course of the infection (e.g. duration), the HPV type(s), the abundance and taxonomic diversity of bacteria, fungi and viruses in the vaginal microbiota, human genetics (SNPs) and basal immunological status.

Technical procedures

DNA extraction

DNA extraction from cervical smears will be performed using Nuclisens EasyMAg from Biomerieux or an equivalent protocol. For the microbiota analyses, special kits involving physical (via beads) and/or enzymatic breaking of the cellular barrier will be favoured following standard protocols to study the vaginal microbiome [36], e.g. the MagAttract® PowerMicrobiome® DNA/RNA kit from Qiagen.

HPV detection, typing and quantification

The participants' infection status (HPV+ or HPV-) will be assessed using the DEIA test, which is based on a PCR of the short SPF10 amplicon [37] and detects all *Alphapapillomaviruses* with great sensitivity.

For peer review only

If the DEIA test is positive, HPVs will be typed using the LiPA25 kit, which is based on the same SPF10-PCR, and has a lower detection threshold compared to other hybridisation-based typing methods [38].

The reason for basing the detection on the DEIA rather than the LiPA25 is that some *Alphapapillomavirus* may be detected by DEIA but not genotyped by LiPA and also that the DEIA is more sensitive than the LiPA. If the DEIA is positive and the LiPA25 is negative, typing will be performed by sequencing the product of a PGMY09/11 PCR [39], which targets another region of the HPV genome than the SPF10 PCR.

The quantification of HPV DNA genome copy number in the samples will be performed using the protocol set up by Micalessi et al. [40].

Cytokine titration

Cytokines can be used as markers of immune activation or immunosuppression and can also inform us on which components of the immune system are involved. Cervical sponges are centrifuged after the addition of 175*μ*L of PBS. Cervical secretions are analysed for a set of 5 to 6 cytokines levels using the Meso Scale Discovery (MSD) Multiplex ELISA platform, which has a low detection threshold and a slowly saturating dose-response curve. Based on earlier results [41, 42], we will first investigate a large panel of 20 cyctokines (IFN-*α*2a, IFN-*γ*, IL-1*α*, IL-5, IL-6, IL-8, IL-10, IL-12, IL-15, IL-17, IL-18, IL-23, IL-25, IP-10, MCP-1, MIP-1*α*, MIP-3*α*, MIP-3*β*, TNF-*α*, TNF-*β*) to choose the most relevant ones for a longitudinal follow-up.

Flow cytometry

Analysing immune cells via flow cytometry is extremely challenging on cells as fragile as the ones from cervical smears. However, several studies suggest that this is feasible [41–43]. Here, we follow the protocol described in [44].

Staining is performed using a Duraclone custom mix targeting CD45, CD3, CD4, CD8, CD16, CD56, CD69, CD161 and TCR*γδ*. The last marker, Live&Dead tests for cellular viability.

Samples are acquired using a Navios flow cytometer (Beckman Coulter, three-laser

configuration).

Sequencing

Sequencing will be performed for microbiota profiling. It involves PCR amplification of the V3-V4 region of 16S RNA for bacteria [45] and ITS1 for fungi [46]. We anticipate that the bacteria should belong to the operational taxonomic units (OTU) described in the five community state types found in vaginal communities [47, 48]. The virome will also be explored using shotgun sequencing and rolling circle PCR amplification [49]. Human genetics are explored using chip sequencing for SNPs.

Statistical analyses

Times series analyses

The core results of the study will come from the longitudinal follow-up of infected women, which will generate time series, i.e. a set of values collected from the same individual over time (Figure 2). There will be several time series per individual (virus load, number of immune cells, cytokine and antibody levels). These time series will be used to fit mathematical viral kinetics models that describe the interaction between viruses, host target cells (here, in the case of HPV, keratinocytes) and the immune response. These models are commonly developed for viral infections [13, 50–52], including those caused by HPVs [53]. We anticipate our follow-up to yield adequate data for such a fit based on the estimated duration of HPV infections (9 to 18 months [5, 7–9]). Furthermore, the weekly self-samples allow us to increase the resolution if necessary.

We will use non-linear mixed effect models [54] to jointly analyse time series from all participants. More precisely, we will rely on *R* packages such as nlme [55] or lme4 [55]. Note that, in addition to estimating model parameters (e.g. life-expectancy of infected cells or virion production rate of infected cells), this approach also allows us to compare biological models using statistical tools based on model likelihood such as Akaike Information Criterion. For an example of such analysis in the case of HIV, see [51].

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Microbiota dynamics

The composition of the vaginal microbiota has already been described and shown to exhibit considerably less diversity than the gut microbiota [47]. The dynamics of this microbiota has also been studied and shown to closely follow menstrual cycles [48].

Using the time series of OTU abundances (measured via 16S RNA sequencing and qPCR) we will infer interaction parameters by assuming an underlying Lotka-Volterra competition model [56]. This work will include time series analysis techniques (e.g. auto-correlation or local similarity analysis) and statistical inference methods in order to infer community structure and interactions from the next-generation sequencing (NGS) datasets [57]. Finally, statistical methods from ecology will also be used to study community diversity (e.g diversity indices) and community assembly, such as cluster and ordination analyses [58].

Genome Wide Association Studies

We will use human single nucleotide polymorphisms (SNPs) inferred by chip sequencing to look for genetic determinants of key traits (e.g. microbiota composition or HPV infection duration). This is classically done by performing a Genome Wide Association Study (GWAS), which is a complex regression method designed for situations where there are many explanatory variables (here millions of SNPs) for a single trait of interest. GWAS will be performed using classical methods [59]. Earlier GWAS studies have been applied to HPV infections for instance to test for determinants to the ability to seroconvert following infection [60] and cervical cancer (see [61] for a review). Here, our expected sample (N = 300 women) is limited but SNPs with large effects have been detected by studies with comparable sizes [62].

Additional analyses

For all collected variables, descriptive statistics will be calculated according to the level of measurement. For metric variables these statistics can be mean and standard deviation as well as quantiles and more robust statistics [63]. In case of categorical variables group proportions and contingency tables are prepared.

Univariate inferential statistics follow a descriptive analysis. Generally, parametric testing

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

procedures are preferred to non-parametric tests, as the former have higher power. That is why, for metric variables, we will first check whether the data can be assumed to be normally distributed. For normally distributed variables, ANOVA statistics are done to detect differences between groups. In case of significance, post-hoc analysis (Tukey test) are planned to reveal pairwise differences. If the data are not normally distributed or ordinally scaled, non-parametric analyses will be used. These contain the Kruskal-Wallis test and the Wilcoxon test as a post-hoc test with an appropriate correction of the significance level. Since the cell counts are expected to be small, Fisher's exact test will be performed for contingency tables instead of the asymptotic χ^2 test for categorical variables.

Sample size calculation

The study will enrol a total of N = 300 women, with N = 150 in a longitudinal study and N = 150 in a cross-sectional study. The goal of the longitudinal study is to follow 75 women longitudinally, preferentially before they are infected (see above). For the following calculations, we assumed a high percentage of lost during follow-up (30%).

With 150 enrolments and considering that the prevalence of HPV infection in young women is ≈ 60% (based on our preliminary data) and 30% of lost to follow-up, we expect to detect (and successfully follow) 63 infections at inclusion [CI95: 51−75, assuming a binomial distribution to calculate the 95% confidence interval].

Among women who are uninfected at the first visit and considering the yearly incidence being close to 30% [64], we expect 12 [CI95: 6−20] to be infected during the first year of follow-up (still assuming 30% of lost to follow-up).

In the end, with 150 enrolments and assuming a high percentage of lost to follow-up (30%), we expect to successfully follow 75 [CI95: 56−95] women infected at different stages of HPV

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

infection: beginning, during and end.

This will be made possible by the probability of transmission of HPV, which is estimated to be ≈ 90% without condom use and still high with condom use (≈ 40%) [34].

Finally, regarding potential interference with the HPV vaccines, we do not anticipate any significant problem for two reasons. First, as mentioned above, the vaccine coverage is low in France [32]. Second, and more importantly, the vaccines only target few HPV types, thus leaving open the possibility of infection by dozens of types. Furthermore, studying the kinetics of a non-vaccine HPV type in a vaccinated woman will be extremely informative, e.g. to detect any potential cross-reactivity [65].

To run cross-sectional analyses (especially on the microbiota and human genetics), we will enrol *N* = 150 additional women who will only perform the inclusion and the results visits. This sample size was chosen to reach that of earlier GWAS studies [61, 62].

Trial governance

Sponsor

This study is sponsored by the Centre Hospitalier Universitaire (CHU) of Montpellier. The CHU is involved in the implementation of the trial, legal/ethical submissions (see below for details on Ethics approval) and implementing the clinical database (eCRF), which is hosted by Ennov-Clinical (ClinSight). The CHU is not involved in the analysis or interpretation of the data. The CHU of Montpellier performs regular quality control assessments. A clinical research assistant will visit the CeGIDD every 4 months to ensure that implementation is in accordance with the protocol. The CHU has taken out insurance from the Société hospitalière d'assurances mutuelles, 18, rue Edouard Rochet-6 9372 Lyon cedex 08 (contract number 138983) through the full research period, covering its own civil liability and that of any agent (clinical or research staff), in accordance with article L.1121-10 of the French Public Health Code.

For peer review only

Scientific committee

The scientific committee comprises the study investigators, clinicians, scientific experts and representatives of the sponsor. The committee meets yearly and is responsible for following research progress, monitoring compliance with good clinical practices and patient safety. It can also decide relevant modification of the protocol. Requests from third parties to access data collected during the study will be evaluated by the committee.

Monitoring

Monitoring is performed during the whole study at CeGIDD according to the sponsor specific SOP. Routine monitoring visits are made by the monitors designated by the sponsor to check compliance with the protocol, the completeness, accuracy and consistency of the data, and adherence to GCP. The principal investigator ensures that eCRFs are completed in a timely manner and must allow periodical access to eCRFs, patient records, drug logs, and all other study-related documents and materials. The frequency of monitoring visits is determined by factors such as study design and the site enrolment requirements but visits will normally occur at least once every 4 months.

Trial registration

The trial has been registered to ClinicalTrials.gov on 27 Oct 2016 with ID number NCT02946346.

Ethics and Dissemination

The PAPCLEAR trial obtained favourable opinions from the Comité de Protection des Personnes (CPP) Sud Méditerranée I on May 11, 2016 (CPP number 16 42, reference number ID RCB 2016-A00712-49); from the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (CCTIRS) on July 12, 2016 (reference number 16.504); and from the Commission Nationale Informatique et Libertés (CNIL) on Dec 16, 2016 (reference number MMS/ABD/AR1612278, decision number DR-2016-488). This trial

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

was authorised by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM) on July 20, 2016 (reference 20160072000007).

The protocol has been modified since its initial version and the latest modification was accepted by the CPP on Dec 12, 2018.

All participants in the study will sign an informed consent form prior to participation.

The results will be published on preprint servers (e.g. BioRxiv), peer-reviewed journals, post-print servers (e.g. HAL) and disseminated through conferences.

Discussion

Expected results

Acute infections by HPVs are important to study because vaccination is most effective when performed before the first infection. However, we currently know very little about the early stages of HPV infections. This clinical study will give us an unprecedented level of detail into the natural history of HPV infections in young women. Variations in virus load over time have been studied but in the context of cervical cancer in older women [66]. In addition, we will also describe the nature and the dynamics of the immune response (local immune cells and cytokines, circulating anti-HPV antibodies) and of the vaginal microbiota. Beyond these kinetics, we will also have access to data such as infection clearance or not in 24 months, presence of more than one HPV type or coinfection by other STIs.

These data will be analysed in the light of numerous cofactors. One of the most important will be human genetics, with the sequencing of millions of SNPs. Others will be related to the sexual behaviour (number of partners, contraception methods, sexual practices) and general lifestyle. We, therefore, expect broader insights regarding sexual health in young women.

For peer review only

Practical and operational issues

One of the main practical challenges resides in the analysis of cervical smears by flow cytometry. Indeed, the tissues are known to be fragile, adhesive and auto-fluorescent. Even though standard protocols now exist [44], they require processing fresh samples in less than 2 hours.

Another potential issue has to do with contamination by HPV DNA between samples, which are frequent in the HPV field due to the robustness of the virions and the sensitivity of the tests. To certify our ability to control for these, we have entered the 2017 GLOBAL HPV DNA Proficiency Panel from the WHO HPV LabNet [67].

Regarding the enrolment of the participants, we do not expect issues with enrolling 150 women in 28 months for the longitudinal study and 150 for the cross-sectional study. This is due to the number of visitors of the centre who fit the inclusion criteria (more than 3,000 per year) and because of earlier high participation rates in the same population ([33] enrolled 1381 participants in 5 months for their study).

Concerning the follow-up, the high incidence rate of HPV can also lead to transient carriage, i.e. women who are positive for a type only at a single visit. This has been observed for instance in longitudinal studies with a tight follow-up interval [21]. To control for this, we will run the HPV detection test on the cells from the cervical smear after washing with RPMI.

For peer review only

Table 1: Summary of the visit schedules and samples take. The cross-sectional study only includes the first two columns (V1 and V2). The ¤indicate samples taken at visits. + participants infected by a HR-HPV for 12 month will have one PBS smear replaced by a Thinprep® smear to perform a cytology and check for lesions. ⊲ this sample is only taken at the first HPV+ visit of a formerly HPV- participant. ⋆ STI detection will be performed at inclusion unless the participant has been tested within the last 3 months and during the study every 6 months if a new partner has been reported or upon request.

For peer review only

	Inclusion ($V1$)	Results ($V2$)	Return (Vi, with $i > 2$)	
Participants	all	all	HPV+	HPV-
Time	day 0	+ 4 weeks	+ 8 weeks	+ 16 weeks
Eligibility	¤			
Consent	¤			
Gynecological consult	¤	¤	¤	¤
Vaginal pH coton swab	¤	¤	¤	¤
2 vaginal swab samples (Copan ESwabTM)	¤	¤	¤	¤
1 ophtalmological sponge sample	¤		¤	¤
1 cervical smear in Thinprep® (cytology)	¤		+	
1 cervical smear in PBS		¤	+	¤
Blood sampling (HPV antibodies)	¤		¤	
Blood sampling (sequencing)	¤			
Blood sampling (immunophenotyping)	¤		◁	
Other STI detection	⋆	⋆	⋆	⋆
Questionnaire #1 (inclusion)	¤			
Questionnaire #2 (visit)		¤	¤	¤
Questionnaire #3 (home)	¤	¤	¤	¤
Returning self-sampling samples		¤	¤	¤
Serious Adverse Event collection		¤	¤	¤

Abbreviations

ANOVA: Analysis of variance,

ASC-US: Atypical squamous cells of undetermined significance,

CD: Cluster of differentiation,

CI95: 95% Confidence interval,

CeGIDD: Centre Gratuit d'Information de Dépistage et de Diagnostic,

CHU: Centre Hospitalier Universitaire,

CIN: Cervical intraepithelial Neoplasia,

CRT: Clinical Research Technician,

ELISA: enzyme-linked immunosorbent assay,

GWAS: Genome Wide Association Study,

HIV: Human Immunodeficiency Virus,

HPV: Human Papillomavirus,

HR: high-risk,

ITS: Internal Transcribed Spacer,

HSIL: High grade Squamous Intraepithelial Lesion,

LR: low-risk,

LSIL: Low grade Squamous Intraepithelial Lesion,

NGS: Next Generation Sequencing,

OTU: Operational Taxonomic Unit,

PBMC: Peripheric Blood Mononuclear Cell,

PBS: Phosphate Buffered Saline,

RPMI: Roswell Park Memorial Institute medium,

SNP: Single Nucleotide Polymorphism,

TCR: T-cell receptor,

WHO: World Health Organisation.

For peer review only

Trial status

The study began on Oct 1, 2016 and the first inclusion was on Nov 3, 2016. On Jun 23, 2018, 89 participants have been included in the longitudinal study. Inclusions in the longitudinal study will continue until March 2019 and the study is expected to last until Aug 2021.

Conflicts of interests

The authors have read and understood BMJ policy on declaration of interests and declare that they have no competing interests.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 648963). The authors acknowledge further support from the Centre National de la Recherche Scientifique (CNRS), the Institute de Recherche pour le Développement (IRD) and the Centre Hospitalier Universitaire (CHU) of Montpellier.

Data statement

All personal and identifying information collected from participants are kept in a secure place at the CeGIDD during the duration of the trial and will be destroyed at the end of the study. The final raw dataset will be accessible only to the sponsor (CHU) and the chief scientist's (SA) team. Anonymous data will be available to external parties upon approval of both the sponsor and the scientific committee. All publications will be made green or gold open access and the corresponding data will be provided as supplementary material or via a public repository (e.g. Zenodo), provided that there is no conflict with ethical guidelines.

Author contributions

Samuel Alizon, Carmen Lia Murall and Massical Rahmoun were the major contributors in the conception of the protocol. Samuel Alizon wrote the initial version of the manuscript. Christian

For peer review only

Selinger, Monique Baldellou, Claire Bernat, Marine Bonneau, Vanina Boué, Mathilde Buisson, Guillaume Christophe, Giuseppe D'Auria, Florence De Taroni, Vincent Foulongne, Rémy Froissart, Christelle Graf, Sophie Grasset, Soraya Groc, Christophe Hirtz, Audrey Jaussent, Julie Lajoie, Frédérique Lorcy, Eric Picot, Marie-Christine Picot, Jacques Ravel, Jacques Reynes, Thérèse Rousset, Aziza Seddiki, Martine Teirlinck, Vincent Tribout, Édouard Tuaillon, Tim Waterboer, Nathalie Jacobs, Ignacio G Bravo, Michel Segondy and Natalie Boulle were involved in the conception of the protocol, in the implementation of the study and read and approved the final manuscript.

For peer review only

Acknowledgements

We thank all the study participants and the CeGIDD staff for their commitment to the study. We also thank the reviewers and, in particular, Dr. Andrew Brouwer for his meticulous reading of the manuscript.

References

1. Tota, J.E., Chevarie-Davis, M., Richardson, L.A., Devries, M., Franco, E.L.: Epidemiology and burden of HPV infection and related diseases: implications for prevention strategies. Prev Med **53 Suppl 1**, 12–21 (2011). doi:10.1016/j.ypmed.2011.08.017

2. Monsonego, J., Zerat, L., Syrjänen, K., Zerat, J.C., Smith, J.S., Halfon, P.: Prevalence of genotype-specific HPV infection among women in France: implications for screening and vaccination. Gynecol Obstet Fertil **41**(5), 305–313 (2013). doi:10.1016/j.gyobfe.2013.03.003

3. Brun-Micaleff, E., Coffy, A., Rey, V., Didelot, M.-N., Combecal, J., Doutre, S., Daurès, J.-P., Segondy, M., Boulle, N.: Cervical cancer screening by cytology and human papillomavirus testing during pregnancy in french women with poor adhesion to regular cervical screening. J Med Virol **86**(3), 536–45 (2014). doi:10.1002/jmv.23764

4. Bruni, L., Diaz, M., Castellsagué, X., Ferrer, E., Bosch, F.X., de Sanjosé, S.: Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. J Infect Dis **202**(12), 1789–99 (2010). doi:10.1086/657321

5. Insinga, R.P., Dasbach, E.J., Elbasha, E.H., Liaw, K.-L., Barr, E.: Incidence and duration of cervical human papillomavirus 6, 11, 16, and 18 infections in young women: an evaluation from multiple analytic perspectives. Cancer Epidemiol Biomarkers Prev **16**(4), 709–15 (2007). doi:10.1158/1055-9965.EPI-06-0846

6. Woodman, C.B.J., Collins, S.I., Young, L.S.: The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer **7**(1), 11–22 (2007). doi:10.1038/nrc2050

7. Rodríguez, A.C., Schiffman, M., Herrero, R., Wacholder, S., Hildesheim, A., Castle, P.E., Solomon, D., Burk, R., Proyecto Epidemiológico Guanacaste Group: Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections. J Natl Cancer Inst **100**(7), 513–7 (2008). doi:10.1093/jnci/djn044

For peer review only

8. Trottier, H., Mahmud, S., Prado, J.C.M., Sobrinho, J.S., Costa, M.C., Rohan, T.E., Villa, L.L., Franco, E.L.: Type-Specific Duration of Human Papillomavirus Infection: Implications for Human Papillomavirus Screening and Vaccination. J Infect Dis **197**(10), 1436–1447 (2008). doi:10.1086/587698

9. Ramanakumar, A.V., Naud, P., Roteli-Martins, C.M., de Carvalho, N.S., de Borba, P.C., Teixeira, J.C., Blatter, M., Moscicki, A.-B., Harper, D.M., Romanowski, B., Tyring, S.K., Ramjattan, B., Schuind, A., Dubin, G., Franco, E.L.: Incidence and duration of type-specific human papillomavirus infection in high-risk HPV-naïve women: results from the control arm of a phase II HPV-16/18 vaccine trial. BMJ Open **6**(8), 011371 (2016). doi:10.1136/bmjopen-2016-011371

10. Houlihan, C.F., Baisley, K., Bravo, I.G., Kapiga, S., de Sanjosé, S., Changalucha, J., Ross, D.A., Hayes, R.J., Watson-Jones, D.: Rapid acquisition of HPV around the time of sexual debut in adolescent girls in Tanzania. Int J Epidemiol **45**(3), 762–773 (2016). doi:10.1093/ije/dyv367

11. Alizon, S., Murall, C.L., Bravo, I.G.: Why Human Papillomavirus Acute Infections Matter. Viruses **9**(10), 293 (2017). doi:10.3390/v9100293

12. Herrero, R., Wacholder, S., Rodríguez, A.C., Solomon, D., González, P., Kreimer, A.R., Porras, C., Schussler, J., Jiménez, S., Sherman, M.E., Quint, W., Schiller, J.T., Lowy, D.R., Schiffman, M., Hildesheim, A., Costa Rica Vaccine Trial Group: Prevention of persistent human papillomavirus infection by an HPV16/18 vaccine: a community-based randomized clinical trial in Guanacaste, Costa Rica. Cancer Discov **1**(5), 408–19 (2011). doi:10.1158/2159-8290.CD-11-0131

13. Canini, L., Perelson, A.S.: Viral kinetic modeling: state of the art. J Pharmacokinet Pharmacodyn **41**(5), 431–443 (2014). doi:10.1007/s10928-014-9363-3

14. Stanley, M.: Immune responses to human papillomavirus. Vaccine **24**(S1), 16–22 (2006).

For peer review only

doi:10.1016/j.vaccine.2005.09.002

15. Ferenczy, A., Franco, E.: Persistent human papillomavirus infection and cervical neoplasia. Lancet Oncol **3**(1), 11–6 (2002)

16. zur Hausen, H.: Review: Papillomaviruses — to Vaccination and Beyond. Biochemistry **73**(5), 498–503 (2008). doi:10.1134/S0006297908050027

17. Einstein, M.H., Schiller, J.T., Viscidi, R.P., Strickler, H.D., Coursaget, P., Tan, T., Halsey, N., Jenkins, D.: Clinician's guide to human papillomavirus immunology: knowns and unknowns. Lancet Infect Dis **9**(6), 347–56 (2009). doi:10.1016/S1473-3099(09)70108-2

18. Van Hede, D., Langers, I., Delvenne, P., Jacobs, N.: Origin and immunoescape of uterine cervical cancer. Presse Med **43**(12P2), 413–421 (2014). doi:10.1016/j.lpm.2014.09.005

19. Stanley, M.: Immunology of HPV infection. Curr Obstet Gynecol Rep **4**(4), 195–200 (2015). doi:10.1007/s13669-015-0134-y. Accessed 2017-03-20

20. Gao, W., Weng, J., Gao, Y., Chen, X.: Comparison of the vaginal microbiota diversity of women with and without human papillomavirus infection: a cross-sectional study. BMC Infect Dis **13**, 271 (2013). doi:10.1186/1471-2334-13-271

21. Brotman, R.M., Shardell, M.D., Gajer, P., Tracy, J.K., Zenilman, J.M., Ravel, J., Gravitt, P.E.: Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. J Infect Dis **210**(11), 1723–33 (2014). doi:10.1093/infdis/jiu330

22. Koutsky, L.A., Ault, K.A., Wheeler, C.M., Brown, D.R., Barr, E., Alvarez, F.B., Chiacchierini, L.M., Jansen, K.U., Proof of Principle Study Investigators: A controlled trial of a human papillomavirus type 16 vaccine. N Engl J Med **347**(21), 1645–51 (2002). doi:10.1056/NEJMoa020586

23. Riethmuller, D., Jacquard, A.-C., Lacau St Guily, J., Aubin, F., Carcopino, X., Pradat, P., Dahlab, A., Prétet, J.-L.: Potential impact of a nonavalent hpv vaccine on the occurrence of hpv-

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

related diseases in france. BMC Public Health **15**, 453 (2015). doi:10.1186/s12889-015-1779-1

24. Joura, E.A., Giuliano, A.R., Iversen, O.-E., Bouchard, C., Mao, C., Mehlsen, J., Moreira, E.D. Jr, Ngan, Y., Petersen, L.K., Lazcano-Ponce, E., Pitisuttithum, P., Restrepo, J.A., Stuart, G., Woelber, L., Yang, Y.C., Cuzick, J., Garland, S.M., Huh, W., Kjaer, S.K., Bautista, O.M., Chan, I.S.F., Chen, J., Gesser, R., Moeller, E., Ritter, M., Vuocolo, S., Luxembourg, A., Broad Spectrum HPV Vaccine Study: A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med **372**(8), 711–23 (2015). doi:10.1056/NEJMoa1405044

25. Murall, C.L., Bauch, C.T., Day, T.: Could the human papillomavirus vaccines drive virulence evolution? Proc Biol Sci **282**, 20141069 (2015). doi:10.1098/rspb.2014.1069

26. Alizon, S., Méthot, P.-O.: Reconciling Pasteur and Darwin to control infectious diseases. PLoS Biol **16**(1), 2003815 (2018). doi:10.1371/journal.pbio.2003815

27. Moscicki, A.-B., Ma, Y., Wibbelsman, C., Darragh, T.M., Powers, A., Farhat, S., Shiboski, S.: Rate of and Risks for Regression of CIN-2 in adolescents and young women. Obstet Gynecol **116**(6), 1373–1380 (2010). doi:10.1097/AOG.0b013e3181fe777f

28. Buck Jr., H.W.: Warts (genital). BMJ Clin Evid **2015**, 1602 (2015)

29. Herrero, R., González, P., Markowitz, L.E.: Present status of human papillomavirus vaccine development and implementation. Lancet Oncol **16**(5), 206–16 (2015). doi:10.1016/S1470-2045(14)70481-4

30. Maver, P.J., Poljak, M.: Progress in prophylactic human papillomavirus (HPV) vaccination in 2016: A literature review. Vaccine (2018). doi:10.1016/j.vaccine.2017.07.113

31. Fagot, J.-P., Boutrelle, A., Ricordeau, P., Weill, A., Allemand, H.: HPV vaccination in France: uptake, costs and issues for the National Health Insurance. Vaccine **29**(19), 3610–6 (2011). doi:10.1016/j.vaccine.2011.02.064

For peer review only

32. Ben Hadj Yahia, M.-B., Dervaux, B., Duport, N., Floret, D., Gaillot, J., Heard, I., Jacquet, A., Le Goaster, C., Lévy-Bruhl, D., Morer, I., Parent du Chatelet, I., Peigue-Lafeuille, H., Rumeau-Pichon, C.: Vaccination contre les infections à papilloamvirus. Technical report, Haut Conseil de la Santé Publique, Paris, France (2014). https://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=454

33. Clarivet, B., Picot, E., Marchandin, H., Tribout, V., Rachedi, N., Schwartzentruber, E., Ledésert, B., Dereure, O., Guillot, B., Picot, M.-C.: Prevalence of Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium in asymptomatic patients under 30 years of age screened in a French sexually transmitted infections clinic. Eur J Dermatol **24**(5), 611–6 (2014). doi:10.1684/ejd.2014.2413

34. Winer, R.L., Hughes, J.P., Feng, Q., O'Reilly, S., Kiviat, N.B., Holmes, K.K., Koutsky, L.A.: Condom use and the risk of genital human papillomavirus infection in young women. N Engl J Med **354**(25), 2645–54 (2006). doi:10.1056/NEJMoa053284

35. Winer, R.L., Hughes, J.P., Feng, Q., Stern, J.E., Xi, L.F., Koutsky, L.A.: Incident Detection of High-Risk Human Papillomavirus Infections in a Cohort of High-Risk Women Aged 25-65 Years. J Infect Dis **214**(5), 665–75 (2016). doi:10.1093/infdis/jiw074

36. Ravel, J., Brotman, R.M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D.W., Sakamoto, J., Koenig, S.S., Fu, L., Zhou, X., Hickey, R.J., Schwebke, J.R., Forney, L.J.: Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. Microbiome **1**(1), 29 (2013). doi:10.1186/2049-2618-1-29

37. Kleter, B., van Doorn, L.-J., ter Schegget, J., Schrauwen, L., van Krimpen, K., Burger, M.,

For peer review only

ter Harmsel, B., Quint, W.: Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses. Am J Pathol **153**(6), 1731–1739 (1998). doi:10.1016/S0002-9440(10)65688-X

38. Geraets, D.T., Struijk, L., Kleter, B., Molijn, A., van Doorn, L.-J., Quint, W.G.V., Colau, B.: The original SPF10 LiPA25 algorithm is more sensitive and suitable for epidemiologic HPV research than the SPF10 INNO-LiPA Extra. J Virol Meth **215-216**, 22–29 (2015). doi:10.1016/j.jviromet.2015.01.001

39. Gravitt, P.E., Peyton, C.L., Alessi, T.Q., Wheeler, C.M., Coutlée, F., Hildesheim, A., Schiffman, M.H., Scott, D.R., Apple, R.J.: Improved amplification of genital Human Papillomaviruses. J Clin Microbiol **38**(1), 357–361 (2000)

40. Micalessi, I.M., Boulet, G.A.V., Bogers, J.J., Benoy, I.H., Depuydt, C.E.: High-throughput detection, genotyping and quantification of the human papillomavirus using real-time PCR. Clin Chem Lab Med **50**(4), 655–61 (2012). doi:10.1515/cclm.2011.835

41. Hunter, P.J., Sheikh, S., David, A.L., Peebles, D.M., Klein, N.: Cervical leukocytes and spontaneous preterm birth. Journal of Reproductive Immunology **113**, 42–49 (2016). doi:10.1016/j.jri.2015.11.002

42. Shannon, B., Yi, T.J., Perusini, S., Gajer, P., Ma, B., Humphrys, M.S., Thomas-Pavanel, J., Chieza, L., Janakiram, P., Saunders, M., Tharao, W., Huibner, S., Shahabi, K., Ravel, J., Rebbapragada, A., Kaul, R.: Association of HPV infection and clearance with cervicovaginal immunology and the vaginal microbiota. Mucosal Immunology **10**(5), 1310–1319 (2017). doi:10.1038/mi.2016.129

43. Lajoie, J., Juno, J., Burgener, A., Rahman, S., Mogk, K., Wachihi, C., Mwanjewe, J., Plummer, F.A., Kimani, J., Ball, T.B., Fowke, K.R.: A distinct cytokine and chemokine profile at the genital mucosa is associated with HIV-1 protection among HIV-exposed seronegative commercial sex workers. Mucosal Immunol **5**(3), 277–287 (2012). doi:10.1038/mi.2012.7

For peer review only

44. Juno, J.A., Boily-Larouche, G., Lajoie, J., Fowke, K.R.: Collection, isolation, and flow cytometric analysis of human endocervical samples. J Vis Exp **89**, 51906 (2014). doi:10.3791/51906

45. Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J.: Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol **74**(8), 2461 (2008). doi:10.1128/AEM.02272-07

46. Findley, K., Oh, J., Yang, J., Conlan, S., Deming, C., Meyer, J.A., Schoenfeld, D., Nomicos, E., Park, M., NIH Intramural Sequencing Center Comparative Sequencing Program, Kong, H.H., Segre, J.A.: Topographic diversity of fungal and bacterial communities in human skin. Nature **498**(7454), 367–370 (2013). doi:10.1038/nature12171. Accessed 2017-09-13

47. Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., Forney, L.J.: Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA **108**, 4680–7 (2011). doi:10.1073/pnas.1002611107

48. Gajer, P., Brotman, R.M., Bai, G., Sakamoto, J., Schütte, U.M.E., Zhong, X., Koenig, S.S.K., Fu, L., Ma, Z.S., Zhou, X., Abdo, Z., Forney, L.J., Ravel, J.: Temporal dynamics of the human vaginal microbiota. Sci Transl Med **4**(132), 132–52 (2012). doi:10.1126/scitranslmed.3003605

49. Johne, R., Müller, H., Rector, A., van Ranst, M., Stevens, H.: Rolling-circle amplification of viral DNA genomes using phi29 polymerase. Trends Microbiol **17**(5), 205–211 (2009). doi:10.1016/j.tim.2009.02.004

50. Nowak, M.A., May, R.M.: Virus Dynamics: Mathematical Principles of Immunology and Virology. Oxford University Press, Oxford, USA (2000)

51. Stafford, M.A., Corey, L., Cao, Y., Daar, E.S., Ho, D.D., Perelson, A.S.: Modeling plasma virus concentration during primary HIV infection. J. theor. Biol. **203**(3), 285–301 (2000).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

doi:10.1006/jtbi.2000.1076

52. Perelson, A.S.: Modelling viral and immune system dynamics. Nat. Rev. Immunol. **2**(1), 28–36 (2002). doi:10.1038/nri700

53. Murall, C.L., Jackson, R., Zehbe, I., Boulle, N., Segondy, M., Alizon, S.: Epithelial stratification shapes infection dynamics. PLoS Comput Biol **15**(1), 1006646 (2019). doi:10.1371/journal.pcbi.1006646

54. Steimer, J.L., Vozeh, S., Racine Poon, A., Holford, N., O'Neil, R.: The population approach: rationale, methods and applications in clinical pharmacology and drug development. In: Balant, P.G.W..L. (ed.) Handbook of Experimental Pharmacology, vol. 110, pp. 405–451. Springer, Berlin (1994)

55. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. Journal of Statistical Software **67**(1) (2015). doi:10.18637/jss.v067.i01

56. Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., Deng, L., Yeliseyev, V., Delaney, M.L., Liu, Q., Olle, B., Stein, R.R., Honda, K., Bry, L., Gerber, G.K.: MDSINE: Microbial dynamical systems INference engine for microbiome time-series analyses. Genome Biology **17**, 121 (2016). doi:10.1186/s13059-016-0980-6. Accessed 2017-03-09

57. Faust, K., Lahti, L., Gonze, D., de Vos, W.M., Raes, J.: Metagenomics meets time series analysis: unraveling microbial community dynamics. Curr Opin Microbiol **25**, 56–66 (2015). doi:10.1016/j.mib.2015.04.004

58. Fox, G.A., Negrete-Yankelevich, S., Sosa, V.J.: Ecological Statistics: Contemporary Theory and Application. Oxford University Press, Oxford, USA (2015)

59. Shi, Y., Li, L., Hu, Z., Li, S., Wang, S., Liu, J., Wu, C., He, L., Zhou, J., Li, Z., Hu, T., Chen, Y., Jia, Y., Wang, S., Wu, L., Cheng, X., Yang, Z., Yang, R., Li, X., Huang, K., Zhang, Q., Zhou, H., Tang, F., Chen, Z., Shen, J., Jiang, J., Ding, H., Xing, H., Zhang, S., Qu, P., Song, X., Lin,

For peer review only

Z., Deng, D., Xi, L., Lv, W., Han, X., Tao, G., Yan, L., Han, Z., Li, Z., Miao, X., Pan, S., Shen, Y., Wang, H., Liu, D., Gong, E., Li, Z., Zhou, L., Luan, X., Wang, C., Song, Q., Wu, S., Xu, H., Shen, J., Qiang, F., Ma, G., Liu, L., Chen, X., Liu, J., Wu, J., Shen, Y., Wen, Y., Chu, M., Yu, J., Hu, X., Fan, Y., He, H., Jiang, Y., Lei, Z., Liu, C., Chen, J., Zhang, Y., Yi, C., Chen, S., Li, W., Wang, D., Wang, Z., Di, W., Shen, K., Lin, D., Shen, H., Feng, Y., Xie, X., Ma, D.: A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. Nat Genet **45**(8), 918–22 (2013). doi:10.1038/ng.2687

60. Chen, D., Gaborieau, V., Zhao, Y., Chabrier, A., Wang, H., Waterboer, T., Zaridze, D., Lissowska, J., Rudnai, P., Fabianova, E., Bencko, V., Janout, V., Foretova, L., Mates, I.N., Szeszenia-Dabrowska, N., Boffetta, P., Pawlita, M., Lathrop, M., Gyllensten, U., Brennan, P., McKay, J.D.: A systematic investigation of the contribution of genetic variation within the MHC region to HPV seropositivity. Hum Mol Genet **24**(9), 2681–2688 (2015). doi:10.1093/hmg/ddv015

61. Chen, D., Gyllensten, U.: Lessons and implications from association studies and post-GWAS analyses of cervical cancer. Trends Genet **31**(1), 41–54 (2015). doi:10.1016/j.tig.2014.10.005

62. Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., Cozzi-Lepri, A., De Luca, A., Easterbrook, P., Francioli, P., Mallal, S., Martinez-Picado, J., Miro, J.M., Obel, N., Smith, J.P., Wyniger, J., Descombes, P., Antonarakis, S.E., Letvin, N.L., McMichael, A.J., Haynes, B.F., Telenti, A., Goldstein, D.B.: A whole-genome association study of major determinants for host control of HIV-1. Science **317**(5840), 944–947 (2007). doi:10.1126/science.1143767

63. Huber, P.J.: The 1972 Wald Lecture Robust Statistics: A Review. Ann Math Stat **43**(4), 1041–1067 (1972). doi:10.1214/aoms/1177692459

64. Winer, R.L., Feng, Q., Hughes, J.P., O'Reilly, S., Kiviat, N.B., Koutsky, L.A.: Risk of female human papillomavirus acquisition associated with first male sex partner. J Infect Dis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

197(2), 279–82 (2008). doi:10.1086/524875

65. Herrero, R.: Human Papillomavirus (HPV) Vaccines: Limited Cross-Protection against Additional HPV Types. J Infect Dis **199**(7), 919–922 (2009). doi:10.1086/597308

66. Depuydt, C.E., Verstraete, L., Berth, M., Beert, J., Bogers, J.-P., Salembier, G., Vereecken, A.J., Bosmans, E.: Human papillomavirus positivity in women undergoing intrauterine insemination has a negative effect on pregnancy rates. Gynecol Obstet Invest **81**(1), 41–6 (2016). doi:10.1159/000434749

67. WHO HPV LabNet. World Health Organization. http://www.who.int/biologicals/areas/human_papillomavirus/WHO_HPV_LabNet/en/

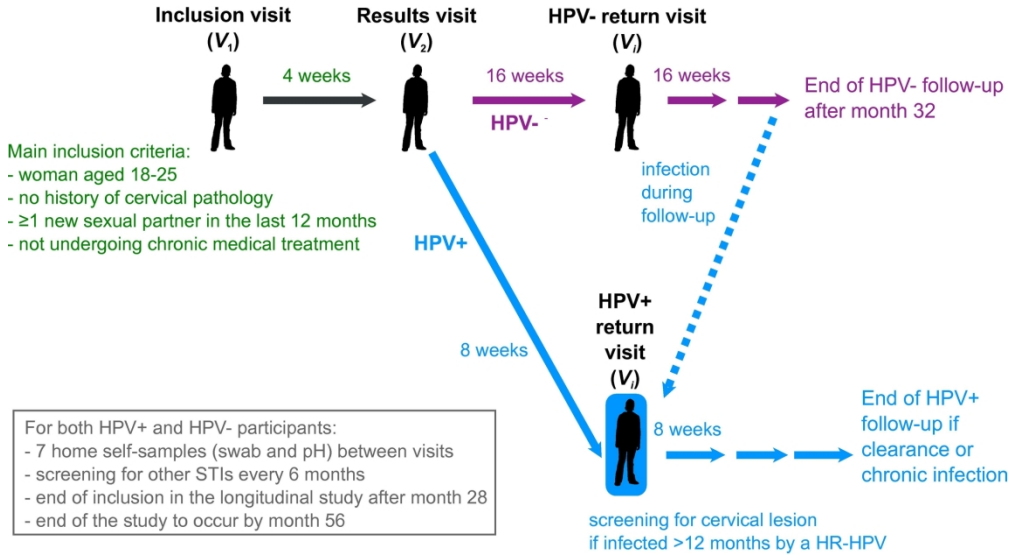
For peer review only

Figure captions

Figure 1: General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V_1), a results visit (V_2) and then return visits (V_i with $i > 2$). For the cross-sectional study, participants only have V_1 and V_2.

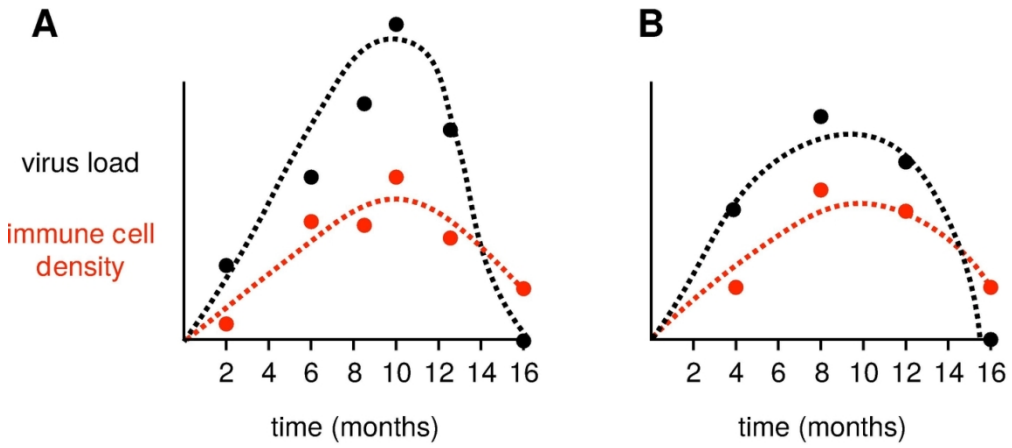
Figure 2: Fitting viral kinetics models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

For peer review only



General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V1), a results visit (V2) and then return visits (Vi with i > 2). For the cross-sectional study, participants only have V1 and V2.

190x104mm (300 x 300 DPI)



Fitting kinetics dynamical models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

120x52mm (300 x 300 DPI)

Note from the Editors: Instructions for reviewers of study protocols

Since launching in 2011, BMJ Open has published study protocols for planned or ongoing research studies. If data collection is complete, we will not consider the manuscript.

Publishing study protocols enables researchers and funding bodies to stay up to date in their fields by providing exposure to research activity that may not otherwise be widely publicised. This can help prevent unnecessary duplication of work and will hopefully enable collaboration. Publishing protocols in full also makes available more information than is currently required by trial registries and increases transparency, making it easier for others (editors, reviewers and readers) to see and understand any deviations from the protocol that occur during the conduct of the study.

The scientific integrity and the credibility of the study data depend substantially on the study design and methodology, which is why the study protocol requires a thorough peer-review.

BMJ Open will consider for publication protocols for any study design, including observational studies and systematic reviews.

Some things to keep in mind when reviewing the study protocol:

- Protocol papers should report planned or ongoing studies. The dates of the study should be included in the manuscript.
- Unfortunately we are unable to customize the reviewer report form for study protocols. As such, some of the items (i.e., those pertaining to results) on the form should be scored as Not Applicable (N/A).
- While some baseline data can be presented, there should be no results or conclusions present in the study protocol.
- For studies that are ongoing, it is generally the case that very few changes can be made to the methodology. As such, requests for revisions are generally clarifications for the rationale or details relating to the methods. If there is a major flaw in the study that would prevent a sound interpretation of the data, we would expect the study protocol to be rejected.

For peer review only

STUDY PROTOCOL

The natural history, dynamics, and ecology of Human papillomaviruses ~~(HPVs)~~ in genital infections of young women: ~~the~~study protocol of the PAPCLEAR cohort study

Carmen Lía Murall[1] , Massilva Rahmoun[1], Christian Selinger[1], Monique Baldellou[2], Claire Bernat[1], Marine Bonneau[3] , Vanina Boué[1] , Mathilde Buisson[4], Guillaume Christophe[2], Giuseppe D'Auria[5,6], Florence De Taroni[2], Vincent Foulongne[7,8], Rémy Froissart[1], Christelle Graf[3], Sophie Grasset[1,2], Soraya Groc[1,7], Christophe Hirtz[9], Audrey Jaussent[10], Julie Lajoie[11], Frédérique Lorcy[12], Eric Picot[2], Marie-Christine Picot[10], Jacques Ravel[13], Jacques Reynes[11], Thérèse Rousset[12], Aziza Seddiki[4], Martine Teirlinck[2], Vincent Tribout[2], Édouard Tuaillon[8], Tim Waterboer[15] , Nathalie Jacobs[16], Ignacio G Bravo[1], Michel Segondy[7,8], Nathalie Boulle[8,12] and Samuel Alizon[1,*]

For peer review only

Word count: 5~~105~~265 words, excluding title page, abstract, references, figures and tables.

1 Laboratoire MIVEGEC (UMR 5290 CNRS, IRD, UM), 911, avenue Agropolis, BP 64501, 34394 Montpellier, France
2 Center for Free Information, Screening and Diagnosis (CeGIDD), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
3 Department of Obstetrics and Gynaecology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France.
4 Department of Research and Innovation (DRI), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
5 Sequencing and Bioinformatics Service, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública), 46020 Valencia, Spain
6 CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain
7 Department of Virology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
8 Pathogenesis and Control of Chronic Infections, INSERM, CHU, University of Montpellier, Montpellier, France
9 University of Montpellier, LBPC/PPC- IRMB, CHU de Montpellier, 80 rue Augustin Fliche, Montpellier, France
10 Department of Medical Information (DIM), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
11 Department of Medical microbiology, University of Manitoba, 745 Bannatyne, Winnipeg, Canada
12 Department of pathology and oncobiology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
13 Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, USA
14 Department of Infectious and Tropical Diseases, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
15 German Cancer Research Center (DKFZ), Infections and Cancer Epidemiology, Im Neuenheimer Feld 280, Heidelberg, Germany
16 GIGA-Research, Cellular and molecular immunology, University of Liège, 3 Avenue de l'Hôpital, 4000 Liège, Belgium

* Author for correspondence: samuel.alizon@cnrs.fr

For peer review only

Abstract

Introduction

Human papillomaviruses (HPVs) are responsible for one third of all cancers caused by infections. Most HPV studies focus on chronic infections and cancers, and thus, we know little about the early stages of viral infection. In particular, the effects of the dynamic interactions between the immune system, the microbiota, and the viral and host genetics on infection clearance or persistence remains poorly understood.

Methods and Analysis

We follow 150 women, aged 18-25 years, longitudinally to monitor immune response features (cytokines and immune cells in the genital tract, circulating anti-HPV antibodies), virus load of HPVs, and vaginal microbiota composition. This is complemented by the assessment of viral and human genetics and behavioural data. To increase the statistical power of the epidemiological arm of the study, an additional 150 women are screened cross-sectionally. ~~This study will provide one of the most detailed follow-up studies of acute HPV infections and their interactions with the host and the vaginal microbiota. It will also allow us to investigate related issues regarding HPV intra-host evolution and diversity, vaginal microbiota dynamics, and sexually transmitted infections.~~

Ethics and Dissemination

This study has been approved by the Comité de Protection des Personnes Sud Méditerranée I (reference number 2016-A00712-49); by the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (reference number 16.504); by the Commission Nationale Informatique et Libertés (reference number MMS/ABD/AR1612278, decision number DR-2016-488) and by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (reference 20160072000007). The results will be published in preprint servers, post-print servers, peer-reviewed journals and disseminated through conferences.

Trial registration number: NCT02946346

Keywords: HPV; acute infection; persistence; virus load; immunity; microbiota; viral kinetics

For peer review only

Article summary

Strengths and limitations of this study

- A ~~D~~dense follow-up with ~~(~~visits every two months for infected women ~~with~~and additional self-sampling every week~~) for N=150 women~~.

- The ~~C~~combination of virological (virus load), immunological (cytokine concentrations and immune cell percentages) and environmental (vaginal microbiota composition, pH) measurements at each visit.

- A limitation is that the density of the follow-up limits the number of participants (N=150), which can affect the power of epidemiological analyses ~~at the epidemiological level~~.

- We complement the longitudinal study with a cross-sectional study of N=150 women to ~~allow for~~ increase ~~the~~ statistical power ~~of epidemiological analyses~~.

For peer review only

Introduction

Epidemiology of HPV genital infections in young adults and public health implications

Infections by Human Papillomaviruses (HPVs) are likely the most common sexually transmitted infection (STI) globally. It is often estimated that, worldwide, more than ~~75~~80% of ~~the~~sexually-active individuals will be infected ~~at some point in their life~~ by an HPV type [1]. In France, a study performed in 2013 in the Paris area estimated the prevalence of HPV genital infections ~~by HPVs at ≈~~ to be 25% ~~of~~in women below 25 years of age [2]. In the area of Montpellier (France), the prevalence of oncogenic HPVs (~~often~~ also referre~~d~~d to as 'high-risk', HR, HPVs) in pregnant women aged 16 to 42 years was close to 20% [3]. These numbers are consistent with worldwide estimates according to which HPVs are most prevalent in women under 25 years of age, with an estimated overall prevalence of 24% [4].

Fortunately, the vast majority of infections by HPVs are asymptomatic and benign. Even for HPV16, which is probably the most oncogenic ~~biologic agent to humans~~human virus, only a minority of infections (less than 10%) become persistent [5], and then a minority of these (12%) progress to cancer if untreated [1, 6]. Indeed, it is estimated that approximately 70 to 100% of infections by HPVs are cleared within 12 to 24 months, with strong differences between virus types [5, 7–9]. Recent studies suggest that primo-infections could be shorter in young girls [10] but, in general, there are many unknowns about the biology of non-persisting infections [11].

Our lack of knowledge partly comes from the fact that in vaccine trials, from which most of the data on infection duration originate, participants are followed every six months for several years [5, 7, 9, 12]. This frequency is sufficient to estimate the time to clearance (or ~~to~~ persistence) but it is not precise enough to understand the within-host dynamics, often referred to as 'kinetics' [13], of infections that last on average 6 to 24 months. Arbitrarily, after 24 months ~~of infection~~, an infection is often considered ~~as being~~to be persistent [14].

Some factors have been shown to correlate with persistence (e.g. immunosuppression,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

smoking, and co-infection with other STIs [15]) but we do not know how these affect viral kinetics. Also, some changes in viral-immunity interactions appear to be related to persistence and disease progression [16–19] but, again, we do not know the underlying interactions between the viruses, the host target cells, and the immune response in acute infections [11]. Finally, it has been argued that the vaginal microbiota may differ between HPV-infected and HPV-uninfected women [20] and that specific microbiota composition may interact with HPV detection [21]. However, it is difficult to disentangle the cause and the consequence. For instance, does the microbiota composition change after the establishment of an HPV infection, or do certain microbiota compositions increase susceptibility to HPV infection?

A better understanding of the within-host infection dynamics and of the determinants of clearance and persistence of viral infection is particularly important in the context of vaccination [22–25]. Indeed, the long-term efficacy of the anti-HPVs vaccines at the population level will largely depend on the within-host viral dynamics because, ultimately, most selective pressures on viral populations occur via the immune response [26]. Furthermore, a better understanding of acute HPV infections can shed ~~a~~ new light on issues related to latency, fertility, or immunotherapies [11].

Prevention strategies and treatment

Treatment

Since most infections by HPVs are benign in young adults and clear within six to 24 months, the current standard of care is to avoid over-treatment, even in the presence of cervical lesions [27]. Clinical interventions (colposcopies, biopsies, and surgery) are less often performed with young women (< 25 years) and only for high-grade (pre-cancerous) lesions (cervical intraepithelial neoplasia grade 2, CIN-2, or more). Low-grade lesions (CIN-1) are not systematically treated but rather monitored yearly to detect any progression to high-grade lesions.

Genital warts caused by non-oncogenic HPVs (often referred to as ‘low-risk’, LR, HPVs) can be removed by surgery or treated with bi- and trichloroacetic acid, cryotherapy or other treatments [28].

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

HPV vaccination

There are currently three licensed vaccines: a bivalent vaccine (Cervarix®) targeting HPV16 and HPV18 (together accounting for 70% of cervical cancers [1]), a quadrivalent vaccine (Gardasil®) that additionally targets HPV6 and HPV11 (non-oncogenic, but highly prevalent and associated to benign proliferative lesions) and, since 2014, a nonavalent vaccine (Gardasil 9®) that targets five more oncogenic types (HPV31, HPV33, HPV45, HPV52, and HPV58, which altogether account for 20% of cervical cancers [24]). These vaccines succeed in eliciting a protective immune response against new infections by the targeted viruses, and are used throughout the world, albeit with wide variation in coverage (for reviews, see e.g. [29, 30]).

Vaccination campaigns in France started in 2006 but with limited coverage: it reached 28.5% in 2008 [31] and has been decreasing ever since [32]. The vaccine is recommended for girls from 11 to 14 years of age, currently with a vaccination scheme of two doses with a six months interval. A catch-up is organised for girls aged 15-19 years, with a three-doses vaccination scheme. Vaccination is reimbursed by the French Social Security but is not mandatory. It is also recommended for men who have sex with men (MSM) as well as for immuno-compromised people [32]. Vaccination is now the primary prevention strategy against cervical cancers.

Screening

In France, the secondary prevention strategy against cervical cancer is routine individual cytology-based screening for pre-cancerous and cancerous cervical lesions in women between 25 and 65 years old. Cytology can also be performed in younger women if they report risk factors for cervical cancer (multiple partners, chronic STIs or HIV infection [32]). Detection of oncogenic HPVs is proposed for triage in case of abnormal cytology (i.e. high-grade or low-grade squamous intraepithelial lesion, HSIL and LSIL respectively, or Atypical Squamous Cells of Undetermined Significance, ASCUS).

Primary objectives

The first primary objective of thise cohort study is to decipher the kinetics and ecology of cervical HPV infections in healthy young women, i.e. follow the population dynamics of the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

virus, the target epithelial cells, and the immune effectors.

The second primary objective is to characterise the diversity of genital HPVs in young women in the region of Montpellier in relationship with their lifestyle, vaccination status, vaginal microbiota, and human genetics.

Secondary objectives

A secondary objective is to characterise the acquisition and clearance dynamics of cervical HPV infections as a function of viral diversity, host immunity, vaginal microbiota and human genetics.

A final objective is to investigate variations in genetic diversity of HPVs during cervical infections.

Methods and analysis

Participants

The study population is composed of young women at risk of HPV infection. The age class was chosen because it exhibits ~~the~~high HPV prevalence ~~of HPV is the highest~~ (24% worldwide [4] and ~~is~~ approximately 25% in France [2]). Inclusion of younger women would have raised technical issues because of the requirement for parental consent.

Women are recruited through a social media page, and through posters and leaflets distributed at the Universities in Montpellier and at the Montpellier STI screening centre (*Centre Gratuit d'Information de Dépistage et de Diagnostic*, CeGIDD). The composition of the population visiting the CeGIDD has already been documented in an earlier study [33]. In total, the centre is visited by approximately 3,000 women per year, the majority of which are under 25 years of age (80%). Approximately 40% of the attendants report three or more partners over the last twelve months and approximately 50% report using adequate behaviour for prevention against HIV.

For peer review only

Inclusion criteria

Participants are women from 18 to 25 years old living in the ~~metropole~~ metropolitan area of Montpellier. They must be sexually active with at least one new partner over the last 12 months. This criteria is fixed to maximise the incidence of new HPV infections. ~~As in any clinical study,~~ ~~p~~Participants must be able to and willing to give written informed consent: they must sign an informed consent form, understand the requirements for the study, and be affiliated to a French social security scheme (which is a state requirement).

Women cannot be included in the study if they have a history of HPV-associated pathology (genital warts or cervical lesions), if they are pregnant or intending to become pregnant in the coming year, infected by HIV, undergoing (or planning to undergo) intense medical treatment (biotherapy, chemotherapy, immunosuppression), planning on moving outside the Montpellier metropolitan area within the next 18 months, in a dependency or employment with the sponsor or the investigator, if they participated in a clinical trial involving administration of drugs within the last four weeks or if they belong to a vulnerable group (e.g. children, adults with physical or mental disabilities).

Design/setting

This study has a longitudinal component aimed at deciphering within-host dynamics and a cross-sectional component~~,~~ aimed at understanding the diversity of HPV infections in young adults in the area of Montpellier, France. The general structure of the study is shown in Fig 1.

If a woman fits the main inclusion criteria, she can go through an inclusion visit (*V*1) with a physician (gynaecologist or midwife) at the CeGIDD. During this visit, ~~she~~the study investigator presents the study and checks all inclusion criteria before asking the participant to read and sign the informed consent form. Participants then undergo a medical consultation during which a number of samples are collected (see below). They then fill out health and lifestyle questionnaires and are given cotton-flocked swabs for self-sampling at home the next visit,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

along with instructions on how to fill in weekly questionnaires through an online form (these are performed throughout the study).

An appointment is scheduled four weeks later for the Results visit (V_2), where the cervical cytology results are communicated. ~~We collect some~~Additional samples are collected and~~and provide more~~ self-sample swabs for home collection are provided.

The next return visits (V_i, where $i > 2$) are as follows:

- Participants with a positive DEIA HPV test (see below), i.e. infected by an *Alphapapillomavirus*, at V_1 join the HPV positive (HPV+) arm of the study with return visits scheduled every 2 months.
- Participants with a negative DEIA HPV test at V_1 join the HPV negative (HPV-) arm with return visits scheduled every 4 months.
- HPV- participants infected by an *Alphapapillomavirus* move to the HPV+ arm.

Intervals between visits are based on earlier results showing that HPV infections last from 9 to 18 months on average depending on the HPV type [5, 7–9] and that a total follow-up of 4 months yields results that are difficult to analyse [21]. The longer interval in the HPV- arm is based on the estimated incidence for HPV genital infections in young women, which is greater than 30% [34, 35].

Participants in the HPV- arm are followed until month 32 of the study.

Participants in the HPV+ arm are followed until they clear the infection or until they have been infected for 24 months (after which we consider that the infection is persistent). Clearance is defined as being negative at two visits in a row for the first HPV type detected in the follow-up.

In between these visits to the CeGIDD, participants are asked to perform regular (every week for HPV+ and every second week for HPV-) self-samples using vaginal swabs, ~~along with a~~to measure ~~of~~ vaginal pH and to fill in~~ing~~ a short questionnaire. Self-samples are stored in the participants' freezer and brought back at every visit.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The study will end with the last HPV+ participant having cleared the infection or been infected for 24 months.

Patients and public involvement

Since all participants are healthy, they are referred to as participants rather than patients. As in any longitudinal study, ensuring participant commitment is challenging. To achieve this goal, we have set up a compensation of 40 EUR per visit and an additional 10 EUR in case of a complete follow-up. Furthermore, participants who have answered a sufficient number of questionnaires and brought back a sufficient number of ~~self-samples~~self-samples ~~will~~ get a 100 EUR bonus at the end. Overall, a participant performing 12 return visits would gain a total compensation of 650 EUR.

Participants did not play a role in the design of this study.

Results of the study will be disseminated to participants who have left the study and to the general public via an email newsletter in French.

Visits

The summary of the visit schedule and of the samples collected at each visit is shown in Table 1.

Inclusion visit (*V*1)

This visit takes place at the CeGIDD and is scheduled by the Clinical Research Technician (~~TEC~~CRT) via phone or email.

Women meet a study investigator, who explains the goals and requirements of the study~~. The physician also~~ and checks that the inclusion criteria are met. If so, after a general discussion, the informed consent forms are signed.

For peer review only

~~The~~A female physician/midwife performs a general exam and then a gynaecological exam during which the following samples are taken:

- vaginal pH cotton swab (EcoCare™),
- vaginal swab (Copan ESwab™) in 1mL Amies liquid for DNA extraction and microbiota analysis,
- vaginal swab (Copan ESwab™) in 1mL of RNA preservation medium,
- ophthalmic sponge (Weck-cel®) to collect cervical secretions for cytokines analysis,
- cervical smear in 20mL of Thinprep® (Preservcyt® liquid) for HPV and HSV assays, and cytology evaluation.

Following the gynaecological consultation, the participant meets with a nurse to measure body temperature, blood pressure and draw 20mL of blood (a 5mL tube for SNPs sequencing, a 10mL tube for immunophenotyping and a 5mL tube for HPV antibody titration). For the longitudinal study, the nurse provides the participant with 3 self-sampling kits, 3 pH strips, a freezer box to bring back to the next visit, as well as instructions on how to perform the home sampling and store the samples in her personal freezer until the next visit.

If the participant has not been tested for a STI in the last 3 months, the nurse draws an additional blood tube of 5mL to test for STIs (HIV, HCV, HBV) and collects vaginal self-samples for chlamydiae and gonorrhea detection. Syphilis testing is prescribed to participants who meet the STI clinic's guidelines.

Finally, the participant meets with the ~~TEC~~CRT to fill in questionnaires #1 (inclusion visit) and #3 (home). The ~~TEC~~CRT answers any remaining questions, explains how to fill the home questionnaires (#3) and sets an appointment for the Results visit.

Results visit (*V*2)

During this visit, the participants are given the result of cervical lesion screening using the liquid cytology (normal, ASCUS, LSIL or HSIL). Participants with a high-grade lesion (HSIL) exit the study and are referred to the gynaecology service of the CHU of Montpellier.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

During this visit, the physician/midwife collects additional samples: 2 vaginal swabs for DNA and RNA analysis, and a cervical smear in 10mL of PBS (to confirm HPV status and perform flow cytometry analyses).

The participant fills in questionnaires #2 (for return visits) and #3 (home). An appointment for the next visit is set and swabs for home self-sampling are given.

Return visits (*Vi*)

These visits only occur in the longitudinal study.

HPV- arm. Participants uninfected by HPV visit the clinic every 4 months until month 26. During these visits, the same samples as in the inclusion visit (*V*1**)** are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

The nurse only draws blood if a screening test for STIs other than HPV is required. The participant then fills in questionnaires #2 and #3 and an appointment is set for the next visit in 16 weeks.

If an HPV infection is detected in the cervical smear collected during this visit, the participant moves to the HPV+ arm and the ~~TEC~~CRT contacts the participant to move her appointment forward.

HPV+ arm. Participants infected by HPV visit the clinic every 2 months. They cannot switch arm and will remain in the HPV+ arm until clearance or the end of the study. During the visits, the same samples as in the inclusion visit (*V*0) are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

The nurse then draws 5mL of blood for HPV antibody titration. If this is the first HPV+ visit following an HPV- visit, the nurse also draws 10mL of blood for immunophenotyping. Finally, if a test for additional STIs is needed, the nurse draws 5mL of blood and collects vaginal self-samples for STI detection.

For peer review only

Importantly, if the participant has been infected by a HR-HPV for more than 12 months and cytology has not been performed within the last 12 months, the cervical smear is put in Thinprep® fixation medium, instead of PBS, for cytological analysis (cervical lesion screening).

Finally, the participant fills in questionnaires #2 and #3, receives self-samples for home collection and an appointment is set for the next visit in 8 weeks.

Endpoints

The primary endpoint for the study is the inclusion and follow-up of HPV-infected women in order to describe the kinetics of HPV virus load, and the associated immune response.

Secondary endpoints are the characterisation of the interactions between the course of the infection (e.g. duration), the HPV type(s), the abundance and taxonomic diversity of bacteria, fungi and viruses in the vaginal microbiota, human genetics (SNPs) and basal immunological status.

Technical procedures

DNA extraction

DNA extraction from cervical smears will be performed using Nuclisens EasyMAg from Biomerieux or an equivalent protocol. For the microbiota analyses, special kits involving physical (via beads) and/or enzymatic breaking of the cellular barrier will be favoured following standard protocols to study the vaginal microbiome [36], e.g. the MagAttract® PowerMicrobiome® DNA/RNA kit from Qiagen.

HPV detection, typing and quantification

The participants' infection status (HPV+ or HPV-) will be assessed using the DEIA test, which is based on a PCR of the short SPF10 amplicon [37] and detects all *Alphapapillomaviruses* with great sensitivity

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

If the DEIA test is positive, HPVs will be typed using the LiPA25 kit, which is based on the same SPF10-PCR, and has a lower detection threshold compared to other hybridisation-based typing methods [38].

The reason for basing the detection on the DEIA rather than the LiPA25 is that some *Alphapapillomavirus* may be detected by DEIA but not genotyped by LiPA and also that the DEIA is more sensitive than the LiPA. If the DEIA is positive and the LiPA25 is negative, typing will be performed by sequencing the product of a PGMY09/11 PCR [39], which targets another region of the HPV genome than the SPF10 PCR.

The quantification of HPV DNA genome copy number in the samples will be performed using the protocol set up by Micalessi et al. [40].

Cytokine titration

Cytokines can be used as markers of immune activation or immunosuppression and can also inform us on which components of the immune system are involved. Cervical sponges are centrifuged after the addition of 175*µ*L of PBS. Cervical secretions are analysed for a set of 5 to 6 cytokines levels using the Meso Scale Discovery (MSD) Multiplex ELISA platform, which has a low detection threshold and a slowly saturating dose-response curve. Based on earlier results [41, 42], we will first investigate a large panel of 20 cyctokines (IFN-*α*2a, IFN-*γ*, IL-1*α*, IL-5, IL-6, IL-8, IL-10, IL-12, IL-15, IL-17, IL-18, IL-23, IL-25, IP-10, MCP-1, MIP-1*α*, MIP-3*α*, MIP-3*β*, TNF-*α*, TNF-*β*) to choose the most relevant ones for a longitudinal follow-up.

Flow cytometry

Analysing immune cells via flow cytometry is extremely challenging on cells as fragile as the ones from cervical smears. However, several studies suggest that this is feasible [41–43]. Here, we follow the protocol described in [44].

Staining~~s are~~ is performed using a Duraclone custom mix targeting CD45, CD3, CD4, CD8, CD16, CD56, CD69, CD161 and TCR*γδ*. The last marker, Live&Dead tests for cellular viability. Samples are acquired using a Navios flow cytometer (Beckman Coulter, three-laser

configuration).

Sequencing

Sequencing will be performed for microbiota profiling. It involves PCR amplification of the V3-V4 region of 16S RNA for bacteria [45] and ITS1 for fungi [46]. We anticipate that the bacteria should belong to the operational taxonomic units (OTU) described in the five community state types found in vaginal communities [47, 48]. The virome will also be explored using shotgun sequencing and rolling circle PCR amplification [49]. Human genetics are explored using chip sequencing for SNPs.

Statistical analyses

Times series analyses

The core results of the study will come from the longitudinal follow-up of infected women, which will generate time series, i.e. a set of values collected from the same individual over time (Figure 2). There will be several time series per individual (virus load, number of immune cells, cytokine and antibody levels). These time series will be used to fit mathematical viral kinetics models that describe the interaction between viruses, host target cells (here, in the case of HPV, keratinocytes) and the immune response. These models are commonly developed for viral infections [13, 50–52], including those caused by HPVs [53]. We anticipate our follow-up to yield adequate data for such a fit based on the estimated duration of HPV infections (9 to 18 months [5, 7–9]). Furthermore, the weekly self-samples allow us to increase the resolution if necessary.

We will use non-linear mixed effect models [54] to jointly analyse time series from all participants. More precisely, we will rely on *R* packages such as nlme [55] or lme4 [55]. Note that, in addition to estimating model parameters (e.g. life-expectancy of infected cells or virion production rate of infected cells), this approach also allows us to compare biological models using statistical tools based on model likelihood such as Akaike Information Criterion. For an example of such analysis in the case of HIV, see [51].

For peer review only

Microbiota dynamics

The composition of the vaginal microbiota has already been described and shown to exhibit considerably less diversity than the gut microbiota [47]. The dynamics of this microbiota has also been studied and shown to closely follow menstrual cycles [48].

Using the time series of OTU abundances (measured via 16S RNA sequencing and qPCR) we will infer interaction parameters by assuming an underlying Lotka-Volterra competition model [56]. This work will include time series analysis techniques (e.g. auto-correlation or local similarity analysis) and statistical inference methods in order to infer community structure and interactions from the next-generation sequencing (NGS) datasets [57]. Finally, statistical methods from ecology will also be used to study community diversity (e.g diversity indices) and community assembly, such as cluster and ordination analyses [58].

Genome Wide Association Studies

We will use human single nucleotide polymorphisms (SNPs) inferred by chip sequencing to look for genetic determinants of key traits (e.g. microbiota composition or HPV infection duration). This is classically done by performing a Genome Wide Association Study (GWAS), which is a complex regression method designed for situations where there are many explanatory variables (here millions of SNPs) for a single trait of interest. GWAS will be performed using classical methods [59]. Earlier GWAS studies have been applied to HPV infections for instance to test for determinants to the ability to seroconvert following infection [60] and cervical cancer (see [61] for a review). Here, our expected sample (N = 300 women) is limited but SNPs with large effects have been detected by studies with comparable sizes [62].

Additional analyses

For all collected variables, descriptive statistics will be calculated according to the level of measurement. For metric variables these statistics can be mean and standard deviation as well as quantiles and more robust statistics [63]. In case of categorical variables group proportions and contingency tables are prepared.

Univariate inferential statistics follow a descriptive analysis. Generally, parametric testing

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

procedures are preferred to non-parametric tests, as the former have higher power. That is why, for metric variables, we will first check whether the data can be assumed to be normally distributed. For normally distributed variables, ANOVA statistics are done to detect differences between groups. In case of significance, post-hoc analysis (Tukey test) are planned to reveal pairwise differences. If the data are not normally distributed or ordinally scaled, non-parametric analyses will be used. These contain the Kruskal-Wallis test and the Wilcoxon test as a post-hoc test with an appropriate correction of the significance level. Since the cell counts are expected to be small, Fisher's exact test will be performed for contingency tables instead of the asymptotic χ^2 test for categorical variables.

Sample size calculation

The study will enrol a total of N = 300 women, with N = 150 in a longitudinal study and N = 150 in a cross-sectional study. The goal of the longitudinal study is to follow 75 women longitudinally, preferentially before they are infected (see above). For the following calculations, we assumed a high percentage of lost during follow-up (30%).

With 150 enrolments and considering that the prevalence of HPV infection in young women is ≈ 60% (based on our preliminary data) and 30% of lost to follow-up, we expect to detect (and successfully follow) 63 infections at inclusion [CI95: 51--75, assuming a binomial distribution to calculate the 95% confidence interval].

Among women who are uninfected at the first visit and considering the yearly incidence being close to 30% [64], we expect 12 [CI95: 6--20] to be infected during the first year of follow-up (still assuming 30% of lost to follow-up).

In the end, with 150 enrolments and assuming a high percentage of lost to follow-up (30%), we expect to successfully follow 75 [CI95: 56--95] women infected at different stages of HPV

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

infection: beginning, during and end.

This will be made possible by the probability of transmission of HPV, which is estimated to be ≈ 90% without condom use and still high with condom use (≈ 40%) [34].

Finally, regarding potential interference with the HPV vaccines, we do not anticipate any significant problem for two reasons. First, as mentioned above, the vaccine coverage is low in France [32]. Second, and more importantly, the vaccines only target few HPV types, thus leaving open the possibility of infection by dozens of types. Furthermore, studying the kinetics of a non-vaccine HPV type in a vaccinated woman will be extremely informative, e.g. to detect any potential cross-reactivity [65].

To run cross-sectional analyses (especially on the microbiota and human genetics), we will enrol *N* = 150 additional women who will only perform the inclusion and the results visits. This sample size was chosen to reach that of earlier GWAS studies [61, 62].

Trial governance

Sponsor

This study is sponsored by the Centre Hospitalier Universitaire (CHU) of Montpellier. The CHU is involved in the implementation of the trial, legal/ethical submissions (see below for details on Ethics approval) and implementing the clinical database (eCRF), which is hosted by Ennov-Clinical (ClinSight). The CHU is not involved in the analysis or interpretation of the data. The CHU of Montpellier performs regular quality control assessments. A clinical research assistant will visit the CeGIDD every 4 months to ensure that implementation is in accordance with the protocol. The CHU has taken out insurance from the Société hospitalière d'assurances mutuelles, 18, rue Edouard Rochet-6 9372 Lyon cedex 08 (contract number 138983) through the full research period, covering its own civil liability and that of any agent (clinical or research staff), in accordance with article L.1121-10 of the French Public Health Code.

For peer review only

Scientific committee

The scientific committee comprises the study investigators, clinicians, scientific experts and representatives of the sponsor. The committee meets yearly and is responsible for following research progress, monitoring compliance with good clinical practices and patient safety. It can also decide relevant modification of the protocol. Requests from third parties to access data collected during the study will be evaluated by the committee.

Monitoring

Monitoring is performed during the whole study at CeGIDD according to the sponsor specific SOP. Routine monitoring visits are made by the monitors designated by the sponsor to check compliance with the protocol, the completeness, accuracy and consistency of the data, and adherence to GCP. The principal investigator ensures that eCRFs are completed in a timely manner and must allow periodical access to eCRFs, patient records, drug logs, and all other study-related documents and materials. The frequency of monitoring visits is determined by factors such as study design and the site enrolment requirements but visits will normally occur at least once every 4 months.

Trial registration

The trial has been registered to ClinicalTrials.gov on 27 Oct 2016 with ID number NCT02946346.

Ethics and Dissemination

The PAPCLEAR trial obtained favourable opinions from the Comité de Protection des Personnes (CPP) Sud Méditerranée I on May 11, 2016 (CPP number 16 42, reference number ID RCB 2016-A00712-49); from the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (CCTIRS) on July 12, 2016 (reference number 16.504); and from the Commission Nationale Informatique et Libertés (CNIL) on Dec 16, 2016 (reference number MMS/ABD/AR1612278, decision number DR-2016-488). This trial

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

was authorised by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM) on July 20, 2016 (reference 20160072000007).

The protocol has been modified since its initial version and the latest modification was accepted by the CPP on Dec 12, 2018.

All participants in the study will sign an informed consent form prior to participation.

The results will be published on preprint servers (e.g. BioRxiv), peer-reviewed journals, post-print servers (e.g. HAL) and disseminated through conferences.

Discussion

Expected results

Acute infections by HPVs are important to study because vaccination is most effective when performed before the first infection. However, we currently know very little about the early stages of HPV infections. This clinical study will give us an unprecedented level of detail into the natural history of HPV infections in young women. Variations in virus load over time have been studied but in the context of cervical cancer in older women [66]. In addition, we will also describe the nature and the dynamics of the immune response (local immune cells and cytokines, circulating anti-HPV antibodies) and of the vaginal microbiota. Beyond these kinetics, we will also have access to data such as infection clearance or not in 24 months, presence of more than one HPV type or coinfection by other STIs.

These data will be analysed in the light of numerous cofactors. One of the most important will be human genetics, with the sequencing of millions of SNPs. Others will be related to the sexual behaviour (number of partners, contraception methods, sexual practices) and general lifestyle. We, therefore, expect broader insights regarding sexual health in young women.

For peer review only

Practical and operational issues

One of the main practical challenges resides in the analysis of cervical smears by flow cytometry. Indeed, the tissues are known to be fragile, adhesive and auto-fluorescent. Even though standard protocols now exist [44], they require processing fresh samples in less than 2 hours.

Another potential issue has to do with contamination by HPV DNA between samples, which are frequent in the HPV field due to the robustness of the virions and the sensitivity of the tests. To certify our ability to control for these, we have entered the 2017 GLOBAL HPV DNA Proficiency Panel from the WHO HPV LabNet [67].

Regarding the enrolment of the participants, we do not expect issues with enrolling 150 women in 28 months for the longitudinal study and 150 for the cross-sectional study. This is due to the number of visitors of the centre who fit the inclusion criteria (more than 3,000 per year) and because of earlier high participation rates in the same population ([33] enrolled 1381 participants in 5 months for their study).

Concerning the follow-up, the high incidence rate of HPV can also lead to transient carriage, i.e. women who are positive for a type only at a single visit. This has been observed for instance in longitudinal studies with a tight follow-up interval [21]. To control for this, we will run the HPV detection test on the cells from the cervical smear after washing with RPMI.

For peer review only

Table 1: Summary of the visit schedules and samples take. The cross-sectional study only includes the first two columns (V1 and V2). The ¤indicate samples taken at visits. + participants infected by a HR-HPV for 12 month will have one PBS smear replaced by a Thinprep® smear to perform a cytology and check for lesions. ◅ this sample is only taken at the first HPV+ visit of a formerly HPV- participant. ⋆ STI detection will be performed at inclusion unless the participant has been tested within the last 3 months and during the study every 6 months if a new partner has been reported or upon request.

For peer review only

	Inclusion ($V1$)	Results ($V2$)	Return (Vi, with $i > 2$)	
Participants	all	all	HPV+	HPV-
Time	day 0	+ 4 weeks	+ 8 weeks	+ 16 weeks
Eligibility	¤			
Consent	¤			
Gynecological consult	¤	¤	¤	¤
Vaginal pH coton swab	¤	¤	¤	¤
2 vaginal swab samples (Copan ESwabTM)	¤	¤	¤	¤
1 ophtalmological sponge sample	¤		¤	¤
1 cervical smear in Thinprep® (cytology)	¤		+	
1 cervical smear in PBS		¤	+	¤
Blood sampling (HPV antibodies)	¤		¤	
Blood sampling (sequencing)	¤			
Blood sampling (immunophenotyping)	¤		◁	
Other STI detection	⋆	⋆	⋆	⋆
Questionnaire #1 (inclusion)	¤			
Questionnaire #2 (visit)		¤	¤	¤
Questionnaire #3 (home)	¤	¤	¤	¤
Returning self-sampling samples		¤	¤	¤
Serious Adverse Event collection		¤	¤	¤

Abbreviations

ANOVA: Analysis of variance,

ASC-US: Atypical squamous cells of undetermined significance,

CD: Cluster of differentiation,

CI95: 95% Confidence interval,

CeGIDD: Centre Gratuit d'Information de Dépistage et de Diagnostic,

CHU: Centre Hospitalier Universitaire,

CIN: Cervical intraepithelial Neoplasia,

CRT: Clinical Research Technician,

ELISA: enzyme-linked immunosorbent assay,

GWAS: Genome Wide Association Study,

HIV: Human Immunodeficiency Virus,

HPV: Human Papillomavirus,

HR: high-risk,

ITS: Internal Transcribed Spacer,

HSIL: High grade Squamous Intraepithelial Lesion,

LR: low-risk,

LSIL: Low grade Squamous Intraepithelial Lesion,

NGS: Next Generation Sequencing,

OTU: Operational Taxonomic Unit,

PBMC: Peripheric Blood Mononuclear Cell,

PBS: Phosphate Buffered Saline,

RPMI: Roswell Park Memorial Institute medium,

SNP: Single Nucleotide Polymorphism,

TCR: T-cell receptor,

WHO: World Health Organisation.

For peer review only

Trial status

The study began on Oct 1, 2016 and the first inclusion was on Nov 3, 2016. On Jun 23, 2018, 89 participants have been included in the longitudinal study. Inclusions in the longitudinal study will continue until March 2019 and the study is expected to last until Aug 2021.

Conflicts of interests

The authors have read and understood BMJ policy on declaration of interests and declare that they have no competing interests.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 648963). The authors acknowledge further support from the Centre National de la Recherche Scientifique (CNRS), the Institute de Recherche pour le Développement (IRD) and the Centre Hospitalier Universitaire (CHU) of Montpellier.

Data statement

All personal and identifying information collected from participants are kept in a secure place at the CeGIDD during the duration of the trial and will be destroyed at the end of the study. The final raw dataset will be accessible only to the sponsor (CHU) and the chief scientist's (SA) team. Anonymous data will be available to external parties upon approval of both the sponsor and the scientific committee. All publications will be made green or gold open access and the corresponding data will be provided as supplementary material or via a public repository (e.g. Zenodo), provided that there is no conflict with ethical guidelines.

Author contributions

Samuel Alizon, Carmen Lia Murall and Massical Rahmoun were the major contributors in the conception of the protocol. Samuel Alizon wrote the initial version of the manuscript. Christian

For peer review only

Selinger, Monique Baldellou, Claire Bernat, Marine Bonneau, Vanina Boué, Mathilde Buisson, Guillaume Christophe, Giuseppe D'Auria, Florence De Taroni, Vincent Foulongne, Rémy Froissart, Christelle Graf, Sophie Grasset, Soraya Groc, Christophe Hirtz, Audrey Jaussent, Julie Lajoie, Frédérique Lorcy, Eric Picot, Marie-Christine Picot, Jacques Ravel, Jacques Reynes, Thérèse Rousset, Aziza Seddiki, Martine Teirlinck, Vincent Tribout, Édouard Tuaillon, Tim Waterboer, Nathalie Jacobs, Ignacio G Bravo, Michel Segondy and Natalie Boulle were involved in the conception of the protocol, in the implementation of the study and read and approved the final manuscript.

~~Ethics approval and consent to participate~~

~~The PAPCLEAR trial obtained favourable opinions from the Comité de Protection des Personnes (CPP) Sud Méditerranée I on May 11, 2016 (CPP number 16 42, reference number ID RCB 2016-A00712-49); from the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (CCTIRS) on July 12, 2016 (reference number 16.504); and from the Commission Nationale Informatique et Libertés (CNIL) on Dec 16, 2016 (reference number MMS/ABD/AR1612278, decision number DR-2016-488). This trial was authorised by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM) on July 20, 2016 (reference 20160072000007).~~

~~The protocol has been modified since its initial version and the latest modification was accepted by the CPP on Dec 12, 2018.~~

~~All participants in the study will sign an informed consent form prior to participation.~~

Acknowledgements

We thank all the study participants and the CeGIDD staff for their commitment to the study. We also thank the reviewers and, in particular, Dr. Andrew Brouwer for his meticulous reading of the manuscript.

For peer review only

References

1. Tota, J.E., Chevarie-Davis, M., Richardson, L.A., Devries, M., Franco, E.L.: Epidemiology and burden of HPV infection and related diseases: implications for prevention strategies. Prev Med **53 Suppl 1**, 12–21 (2011). doi:10.1016/j.ypmed.2011.08.017

2. Monsonego, J., Zerat, L., Syrjänen, K., Zerat, J.C., Smith, J.S., Halfon, P.: Prevalence of genotype-specific HPV infection among women in France: implications for screening and vaccination. Gynecol Obstet Fertil **41**(5), 305–313 (2013). doi:10.1016/j.gyobfe.2013.03.003

3. Brun-Micaleff, E., Coffy, A., Rey, V., Didelot, M.-N., Combecal, J., Doutre, S., Daurès, J.-P., Segondy, M., Boulle, N.: Cervical cancer screening by cytology and human papillomavirus testing during pregnancy in french women with poor adhesion to regular cervical screening. J Med Virol **86**(3), 536–45 (2014). doi:10.1002/jmv.23764

4. Bruni, L., Diaz, M., Castellsagué, X., Ferrer, E., Bosch, F.X., de Sanjosé, S.: Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. J Infect Dis **202**(12), 1789–99 (2010). doi:10.1086/657321

5. Insinga, R.P., Dasbach, E.J., Elbasha, E.H., Liaw, K.-L., Barr, E.: Incidence and duration of cervical human papillomavirus 6, 11, 16, and 18 infections in young women: an evaluation from multiple analytic perspectives. Cancer Epidemiol Biomarkers Prev **16**(4), 709–15 (2007). doi:10.1158/1055-9965.EPI-06-0846

6. Woodman, C.B.J., Collins, S.I., Young, L.S.: The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer **7**(1), 11–22 (2007). doi:10.1038/nrc2050

7. Rodríguez, A.C., Schiffman, M., Herrero, R., Wacholder, S., Hildesheim, A., Castle, P.E., Solomon, D., Burk, R., Proyecto Epidemiológico Guanacaste Group: Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections. J Natl Cancer Inst **100**(7), 513–7 (2008). doi:10.1093/jnci/djn044

For peer review only

8. Trottier, H., Mahmud, S., Prado, J.C.M., Sobrinho, J.S., Costa, M.C., Rohan, T.E., Villa, L.L., Franco, E.L.: Type-Specific Duration of Human Papillomavirus Infection: Implications for Human Papillomavirus Screening and Vaccination. J Infect Dis **197**(10), 1436–1447 (2008). doi:10.1086/587698

9. Ramanakumar, A.V., Naud, P., Roteli-Martins, C.M., de Carvalho, N.S., de Borba, P.C., Teixeira, J.C., Blatter, M., Moscicki, A.-B., Harper, D.M., Romanowski, B., Tyring, S.K., Ramjattan, B., Schuind, A., Dubin, G., Franco, E.L.: Incidence and duration of type-specific human papillomavirus infection in high-risk HPV-naïve women: results from the control arm of a phase II HPV-16/18 vaccine trial. BMJ Open **6**(8), 011371 (2016). doi:10.1136/bmjopen-2016-011371

10. Houlihan, C.F., Baisley, K., Bravo, I.G., Kapiga, S., de Sanjosé, S., Changalucha, J., Ross, D.A., Hayes, R.J., Watson-Jones, D.: Rapid acquisition of HPV around the time of sexual debut in adolescent girls in Tanzania. Int J Epidemiol **45**(3), 762–773 (2016). doi:10.1093/ije/dyv367

11. Alizon, S., Murall, C.L., Bravo, I.G.: Why Human Papillomavirus Acute Infections Matter. Viruses **9**(10), 293 (2017). doi:10.3390/v9100293

12. Herrero, R., Wacholder, S., Rodríguez, A.C., Solomon, D., González, P., Kreimer, A.R., Porras, C., Schussler, J., Jiménez, S., Sherman, M.E., Quint, W., Schiller, J.T., Lowy, D.R., Schiffman, M., Hildesheim, A., Costa Rica Vaccine Trial Group: Prevention of persistent human papillomavirus infection by an HPV16/18 vaccine: a community-based randomized clinical trial in Guanacaste, Costa Rica. Cancer Discov **1**(5), 408–19 (2011). doi:10.1158/2159-8290.CD-11-0131

13. Canini, L., Perelson, A.S.: Viral kinetic modeling: state of the art. J Pharmacokinet Pharmacodyn **41**(5), 431–443 (2014). doi:10.1007/s10928-014-9363-3

14. Stanley, M.: Immune responses to human papillomavirus. Vaccine **24**(S1), 16–22 (2006).

For peer review only

doi:10.1016/j.vaccine.2005.09.002

15. Ferenczy, A., Franco, E.: Persistent human papillomavirus infection and cervical neoplasia. Lancet Oncol **3**(1), 11–6 (2002)

16. zur Hausen, H.: Review: Papillomaviruses — to Vaccination and Beyond. Biochemistry **73**(5), 498–503 (2008). doi:10.1134/S0006297908050027

17. Einstein, M.H., Schiller, J.T., Viscidi, R.P., Strickler, H.D., Coursaget, P., Tan, T., Halsey, N., Jenkins, D.: Clinician's guide to human papillomavirus immunology: knowns and unknowns. Lancet Infect Dis **9**(6), 347–56 (2009). doi:10.1016/S1473-3099(09)70108-2

18. Van Hede, D., Langers, I., Delvenne, P., Jacobs, N.: Origin and immunoescape of uterine cervical cancer. Presse Med **43**(12P2), 413–421 (2014). doi:10.1016/j.lpm.2014.09.005

19. Stanley, M.: Immunology of HPV infection. Curr Obstet Gynecol Rep **4**(4), 195–200 (2015). doi:10.1007/s13669-015-0134-y. Accessed 2017-03-20

20. Gao, W., Weng, J., Gao, Y., Chen, X.: Comparison of the vaginal microbiota diversity of women with and without human papillomavirus infection: a cross-sectional study. BMC Infect Dis **13**, 271 (2013). doi:10.1186/1471-2334-13-271

21. Brotman, R.M., Shardell, M.D., Gajer, P., Tracy, J.K., Zenilman, J.M., Ravel, J., Gravitt, P.E.: Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. J Infect Dis **210**(11), 1723–33 (2014). doi:10.1093/infdis/jiu330

22. Koutsky, L.A., Ault, K.A., Wheeler, C.M., Brown, D.R., Barr, E., Alvarez, F.B., Chiacchierini, L.M., Jansen, K.U., Proof of Principle Study Investigators: A controlled trial of a human papillomavirus type 16 vaccine. N Engl J Med **347**(21), 1645–51 (2002). doi:10.1056/NEJMoa020586

23. Riethmuller, D., Jacquard, A.-C., Lacau St Guily, J., Aubin, F., Carcopino, X., Pradat, P., Dahlab, A., Prétet, J.-L.: Potential impact of a nonavalent hpv vaccine on the occurrence of hpv-

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

related diseases in france. BMC Public Health **15**, 453 (2015). doi:10.1186/s12889-015-1779-1

24. Joura, E.A., Giuliano, A.R., Iversen, O.-E., Bouchard, C., Mao, C., Mehlsen, J., Moreira, E.D. Jr, Ngan, Y., Petersen, L.K., Lazcano-Ponce, E., Pitisuttithum, P., Restrepo, J.A., Stuart, G., Woelber, L., Yang, Y.C., Cuzick, J., Garland, S.M., Huh, W., Kjaer, S.K., Bautista, O.M., Chan, I.S.F., Chen, J., Gesser, R., Moeller, E., Ritter, M., Vuocolo, S., Luxembourg, A., Broad Spectrum HPV Vaccine Study: A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med **372**(8), 711–23 (2015). doi:10.1056/NEJMoa1405044

25. Murall, C.L., Bauch, C.T., Day, T.: Could the human papillomavirus vaccines drive virulence evolution? Proc Biol Sci **282**, 20141069 (2015). doi:10.1098/rspb.2014.1069

26. Alizon, S., Méthot, P.-O.: Reconciling Pasteur and Darwin to control infectious diseases. PLoS Biol **16**(1), 2003815 (2018). doi:10.1371/journal.pbio.2003815

27. Moscicki, A.-B., Ma, Y., Wibbelsman, C., Darragh, T.M., Powers, A., Farhat, S., Shiboski, S.: Rate of and Risks for Regression of CIN-2 in adolescents and young women. Obstet Gynecol **116**(6), 1373–1380 (2010). doi:10.1097/AOG.0b013e3181fe777f

28. Buck Jr., H.W.: Warts (genital). BMJ Clin Evid **2015**, 1602 (2015)

29. Herrero, R., González, P., Markowitz, L.E.: Present status of human papillomavirus vaccine development and implementation. Lancet Oncol **16**(5), 206–16 (2015). doi:10.1016/S1470-2045(14)70481-4

30. Maver, P.J., Poljak, M.: Progress in prophylactic human papillomavirus (HPV) vaccination in 2016: A literature review. Vaccine (2018). doi:10.1016/j.vaccine.2017.07.113

31. Fagot, J.-P., Boutrelle, A., Ricordeau, P., Weill, A., Allemand, H.: HPV vaccination in France: uptake, costs and issues for the National Health Insurance. Vaccine **29**(19), 3610–6 (2011). doi:10.1016/j.vaccine.2011.02.064

For peer review only

32. Ben Hadj Yahia, M.-B., Dervaux, B., Duport, N., Floret, D., Gaillot, J., Heard, I., Jacquet, A., Le Goaster, C., Lévy-Bruhl, D., Morer, I., Parent du Chatelet, I., Peigue-Lafeuille, H., Rumeau-Pichon, C.: Vaccination contre les infections à papilloamvirus. Technical report, Haut Conseil de la Santé Publique, Paris, France (2014). https://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=454

33. Clarivet, B., Picot, E., Marchandin, H., Tribout, V., Rachedi, N., Schwartzentruber, E., Ledésert, B., Dereure, O., Guillot, B., Picot, M.-C.: Prevalence of Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium in asymptomatic patients under 30 years of age screened in a French sexually transmitted infections clinic. Eur J Dermatol **24**(5), 611–6 (2014). doi:10.1684/ejd.2014.2413

34. Winer, R.L., Hughes, J.P., Feng, Q., O'Reilly, S., Kiviat, N.B., Holmes, K.K., Koutsky, L.A.: Condom use and the risk of genital human papillomavirus infection in young women. N Engl J Med **354**(25), 2645–54 (2006). doi:10.1056/NEJMoa053284

35. Winer, R.L., Hughes, J.P., Feng, Q., Stern, J.E., Xi, L.F., Koutsky, L.A.: Incident Detection of High-Risk Human Papillomavirus Infections in a Cohort of High-Risk Women Aged 25-65 Years. J Infect Dis **214**(5), 665–75 (2016). doi:10.1093/infdis/jiw074

36. Ravel, J., Brotman, R.M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D.W., Sakamoto, J., Koenig, S.S., Fu, L., Zhou, X., Hickey, R.J., Schwebke, J.R., Forney, L.J.: Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. Microbiome **1**(1), 29 (2013). doi:10.1186/2049-2618-1-29

37. Kleter, B., van Doorn, L.-J., ter Schegget, J., Schrauwen, L., van Krimpen, K., Burger, M.,

For peer review only

ter Harmsel, B., Quint, W.: Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses. Am J Pathol **153**(6), 1731–1739 (1998). doi:10.1016/S0002-9440(10)65688-X

38. Geraets, D.T., Struijk, L., Kleter, B., Molijn, A., van Doorn, L.-J., Quint, W.G.V., Colau, B.: The original SPF10 LiPA25 algorithm is more sensitive and suitable for epidemiologic HPV research than the SPF10 INNO-LiPA Extra. J Virol Meth **215-216**, 22–29 (2015). doi:10.1016/j.jviromet.2015.01.001

39. Gravitt, P.E., Peyton, C.L., Alessi, T.Q., Wheeler, C.M., Coutlée, F., Hildesheim, A., Schiffman, M.H., Scott, D.R., Apple, R.J.: Improved amplification of genital Human Papillomaviruses. J Clin Microbiol **38**(1), 357–361 (2000)

40. Micalessi, I.M., Boulet, G.A.V., Bogers, J.J., Benoy, I.H., Depuydt, C.E.: High-throughput detection, genotyping and quantification of the human papillomavirus using real-time PCR. Clin Chem Lab Med **50**(4), 655–61 (2012). doi:10.1515/cclm.2011.835

41. Hunter, P.J., Sheikh, S., David, A.L., Peebles, D.M., Klein, N.: Cervical leukocytes and spontaneous preterm birth. Journal of Reproductive Immunology **113**, 42–49 (2016). doi:10.1016/j.jri.2015.11.002

42. Shannon, B., Yi, T.J., Perusini, S., Gajer, P., Ma, B., Humphrys, M.S., Thomas-Pavanel, J., Chieza, L., Janakiram, P., Saunders, M., Tharao, W., Huibner, S., Shahabi, K., Ravel, J., Rebbapragada, A., Kaul, R.: Association of HPV infection and clearance with cervicovaginal immunology and the vaginal microbiota. Mucosal Immunology **10**(5), 1310–1319 (2017). doi:10.1038/mi.2016.129

43. Lajoie, J., Juno, J., Burgener, A., Rahman, S., Mogk, K., Wachihi, C., Mwanjewe, J., Plummer, F.A., Kimani, J., Ball, T.B., Fowke, K.R.: A distinct cytokine and chemokine profile at the genital mucosa is associated with HIV-1 protection among HIV-exposed seronegative commercial sex workers. Mucosal Immunol **5**(3), 277–287 (2012). doi:10.1038/mi.2012.7

For peer review only

44. Juno, J.A., Boily-Larouche, G., Lajoie, J., Fowke, K.R.: Collection, isolation, and flow cytometric analysis of human endocervical samples. J Vis Exp **89**, 51906 (2014). doi:10.3791/51906

45. Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J.: Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol **74**(8), 2461 (2008). doi:10.1128/AEM.02272-07

46. Findley, K., Oh, J., Yang, J., Conlan, S., Deming, C., Meyer, J.A., Schoenfeld, D., Nomicos, E., Park, M., NIH Intramural Sequencing Center Comparative Sequencing Program, Kong, H.H., Segre, J.A.: Topographic diversity of fungal and bacterial communities in human skin. Nature **498**(7454), 367–370 (2013). doi:10.1038/nature12171. Accessed 2017-09-13

47. Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., Forney, L.J.: Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA **108**, 4680–7 (2011). doi:10.1073/pnas.1002611107

48. Gajer, P., Brotman, R.M., Bai, G., Sakamoto, J., Schütte, U.M.E., Zhong, X., Koenig, S.S.K., Fu, L., Ma, Z.S., Zhou, X., Abdo, Z., Forney, L.J., Ravel, J.: Temporal dynamics of the human vaginal microbiota. Sci Transl Med **4**(132), 132–52 (2012). doi:10.1126/scitranslmed.3003605

49. Johne, R., Müller, H., Rector, A., van Ranst, M., Stevens, H.: Rolling-circle amplification of viral DNA genomes using phi29 polymerase. Trends Microbiol **17**(5), 205–211 (2009). doi:10.1016/j.tim.2009.02.004

50. Nowak, M.A., May, R.M.: Virus Dynamics: Mathematical Principles of Immunology and Virology. Oxford University Press, Oxford, USA (2000)

51. Stafford, M.A., Corey, L., Cao, Y., Daar, E.S., Ho, D.D., Perelson, A.S.: Modeling plasma virus concentration during primary HIV infection. J. theor. Biol. **203**(3), 285–301 (2000).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

doi:10.1006/jtbi.2000.1076

52. Perelson, A.S.: Modelling viral and immune system dynamics. Nat. Rev. Immunol. **2**(1), 28–36 (2002). doi:10.1038/nri700

53. Murall, C.L., Jackson, R., Zehbe, I., Boulle, N., Segondy, M., Alizon, S.: Epithelial stratification shapes infection dynamics. PLoS Comput Biol **15**(1), 1006646 (2019). doi:10.1371/journal.pcbi.1006646

54. Steimer, J.L., Vozeh, S., Racine Poon, A., Holford, N., O'Neil, R.: The population approach: rationale, methods and applications in clinical pharmacology and drug development. In: Balant, P.G.W..L. (ed.) Handbook of Experimental Pharmacology, vol. 110, pp. 405–451. Springer, Berlin (1994)

55. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. Journal of Statistical Software **67**(1) (2015). doi:10.18637/jss.v067.i01

56. Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., Deng, L., Yeliseyev, V., Delaney, M.L., Liu, Q., Olle, B., Stein, R.R., Honda, K., Bry, L., Gerber, G.K.: MDSINE: Microbial dynamical systems INference engine for microbiome time-series analyses. Genome Biology **17**, 121 (2016). doi:10.1186/s13059-016-0980-6. Accessed 2017-03-09

57. Faust, K., Lahti, L., Gonze, D., de Vos, W.M., Raes, J.: Metagenomics meets time series analysis: unraveling microbial community dynamics. Curr Opin Microbiol **25**, 56–66 (2015). doi:10.1016/j.mib.2015.04.004

58. Fox, G.A., Negrete-Yankelevich, S., Sosa, V.J.: Ecological Statistics: Contemporary Theory and Application. Oxford University Press, Oxford, USA (2015)

59. Shi, Y., Li, L., Hu, Z., Li, S., Wang, S., Liu, J., Wu, C., He, L., Zhou, J., Li, Z., Hu, T., Chen, Y., Jia, Y., Wang, S., Wu, L., Cheng, X., Yang, Z., Yang, R., Li, X., Huang, K., Zhang, Q., Zhou, H., Tang, F., Chen, Z., Shen, J., Jiang, J., Ding, H., Xing, H., Zhang, S., Qu, P., Song, X., Lin,

For peer review only

Z., Deng, D., Xi, L., Lv, W., Han, X., Tao, G., Yan, L., Han, Z., Li, Z., Miao, X., Pan, S., Shen, Y., Wang, H., Liu, D., Gong, E., Li, Z., Zhou, L., Luan, X., Wang, C., Song, Q., Wu, S., Xu, H., Shen, J., Qiang, F., Ma, G., Liu, L., Chen, X., Liu, J., Wu, J., Shen, Y., Wen, Y., Chu, M., Yu, J., Hu, X., Fan, Y., He, H., Jiang, Y., Lei, Z., Liu, C., Chen, J., Zhang, Y., Yi, C., Chen, S., Li, W., Wang, D., Wang, Z., Di, W., Shen, K., Lin, D., Shen, H., Feng, Y., Xie, X., Ma, D.: A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. Nat Genet **45**(8), 918–22 (2013). doi:10.1038/ng.2687

60. Chen, D., Gaborieau, V., Zhao, Y., Chabrier, A., Wang, H., Waterboer, T., Zaridze, D., Lissowska, J., Rudnai, P., Fabianova, E., Bencko, V., Janout, V., Foretova, L., Mates, I.N., Szeszenia-Dabrowska, N., Boffetta, P., Pawlita, M., Lathrop, M., Gyllensten, U., Brennan, P., McKay, J.D.: A systematic investigation of the contribution of genetic variation within the MHC region to HPV seropositivity. Hum Mol Genet **24**(9), 2681–2688 (2015). doi:10.1093/hmg/ddv015

61. Chen, D., Gyllensten, U.: Lessons and implications from association studies and post-GWAS analyses of cervical cancer. Trends Genet **31**(1), 41–54 (2015). doi:10.1016/j.tig.2014.10.005

62. Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., Cozzi-Lepri, A., De Luca, A., Easterbrook, P., Francioli, P., Mallal, S., Martinez-Picado, J., Miro, J.M., Obel, N., Smith, J.P., Wyniger, J., Descombes, P., Antonarakis, S.E., Letvin, N.L., McMichael, A.J., Haynes, B.F., Telenti, A., Goldstein, D.B.: A whole-genome association study of major determinants for host control of HIV-1. Science **317**(5840), 944–947 (2007). doi:10.1126/science.1143767

63. Huber, P.J.: The 1972 Wald Lecture Robust Statistics: A Review. Ann Math Stat **43**(4), 1041–1067 (1972). doi:10.1214/aoms/1177692459

64. Winer, R.L., Feng, Q., Hughes, J.P., O'Reilly, S., Kiviat, N.B., Koutsky, L.A.: Risk of female human papillomavirus acquisition associated with first male sex partner. J Infect Dis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

197(2), 279–82 (2008). doi:10.1086/524875

65. Herrero, R.: Human Papillomavirus (HPV) Vaccines: Limited Cross-Protection against Additional HPV Types. J Infect Dis **199**(7), 919–922 (2009). doi:10.1086/597308

66. Depuydt, C.E., Verstraete, L., Berth, M., Beert, J., Bogers, J.-P., Salembier, G., Vereecken, A.J., Bosmans, E.: Human papillomavirus positivity in women undergoing intrauterine insemination has a negative effect on pregnancy rates. Gynecol Obstet Invest **81**(1), 41–6 (2016). doi:10.1159/000434749

67. WHO HPV LabNet. World Health Organization. http://www.who.int/biologicals/areas/human_papillomavirus/WHO_HPV_LabNet/en/

For peer review only

Figure captions

Figure 1: General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V_1), a results visit (V_2) and then return visits (V_i with $i > 2$). For the cross-sectional study, participants only have V_1 and V_2.

Figure 2: Fitting viral kinetics models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

For peer review only

BMJ Open

The natural history, dynamics, and ecology of Human papillomaviruses in genital infections of young women: protocol of the PAPCLEAR cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025129.R3
Article Type:	Protocol
Date Submitted by the Author:	16-Apr-2019
Complete List of Authors:	Murall, Carmen Lia; Centre National de la Recherche Scientifique, MIVEGEC Rahmoun, Massilva; Centre National de la Recherche Scientifique, MIVEGEC Selinger, Christian; Centre National de la Recherche Scientifique, MIVEGEC Baldellou, Monique; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Bernat, Claire; Centre National de la Recherche Scientifique, MIVEGEC Bonneau, Marine; Centre Hospitalier Regional Universitaire de Montpellier, Department of Obstetrics and Gynaecology Boué, Vanina; Centre National de la Recherche Scientifique, MIVEGEC Buisson, Mathilde; Centre Hospitalier Regional Universitaire de Montpellier, Department of Research and Innovation Christophe, Guillaume; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD D'Auria, Giuseppe; Fundacio per al Foment de la Investigacio Sanitaria i Biomedica, Sequencing and Bioinformatics Service; Centro de Investigacion Biomedica en Red de Epidemiologia y Salud Publica De Taroni, Florence; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Foulongne, Vincent; Centre Hospitalier Regional Universitaire de Montpellier, Department of Bacteriology and Virology Froissart, Rémy; Centre National de la Recherche Scientifique, MIVEGEC Graf, Christelle; Centre Hospitalier Regional Universitaire de Montpellier, Department of Obstetrics and Gynaecology Grasset, Sophie; Centre Hospitalier Regional Universitaire de Montpellier, Department of Virology; Centre National de la Recherche Scientifique, MIVEGEC Groc, Soraya; Centre Hospitalier Regional Universitaire de Montpellier, Department of Virology; Centre National de la Recherche Scientifique, MIVEGEC Hirtz, Christophe; Centre Hospitalier Regional Universitaire de Montpellier, LBPC/PPC, IRMB Jaussent, Audrey; Centre Hospitalier Regional Universitaire de Montpellier, Department of Medical Information Lajoie, Julie; University of Manitoba College of Medicine, Department of Medical microbiology Lorcy, Frédérique; Centre Hospitalier Regional Universitaire de

	Montpellier, Laboratoire d'anatomie et cytologie pathologiques Picot, Eric; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD PICOT, Marie-Christine; Centre Hospitalier Regional Universitaire de Montpellier, Department of Medical Information Ravel, Jacques; University of Maryland School of Medicine, Institute for Genome Sciences Reynes, Jacques; Centre Hospitalier Regional Universitaire de Montpellier, Department of Infectious and Tropical Diseases Rousset, Thérèse; Centre Hospitalier Regional Universitaire de Montpellier, Department of pathology and oncobiology Seddiki, Aziza; Centre Hospitalier Regional Universitaire de Montpellier, Department of Research and Innovation Teirlinck, Martine; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Tribout, Vincent; Centre Hospitalier Regional Universitaire de Montpellier, CeGIDD Tuaillon, Edouard; Centre Hospitalier Regional Universitaire de Montpellier, Department of bacteriology and virology Waterboer, Tim; Deutsches Krebsforschungszentrum, Infections and Cancer Epidemiology Jacobs, Nathalie; Universite de Liege Faculte des Sciences, GIGA-Research, Cellular and molecular immunology Bravo, Ignacio; Centre National de la Recherche Scientifique, MIVEGEC Segondy, Michel; Centre Hospitalier Regional Universitaire de Montpellier, Department of Bacteriology and Virology Boulle, Nathalie; Centre Hospitalier Regional Universitaire de Montpellier, Department of pathology and oncobiology Alizon, Samuel; Centre National de la Recherche Scientifique, MIVEGEC
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology, Immunology (including allergy), Genetics and genomics
Keywords:	VIROLOGY, IMMUNOLOGY, Epidemiology < INFECTIOUS DISEASES, MICROBIOLOGY, GENETICS

SCHOLARONE™
Manuscripts

STUDY PROTOCOL

The natural history, dynamics, and ecology of Human papillomaviruses in genital infections of young women: protocol of the PAPCLEAR cohort study

Carmen Lía Murall[1] , Massilva Rahmoun[1], Christian Selinger[1], Monique Baldellou[2], Claire Bernat[1], Marine Bonneau[3] , Vanina Boué[1] , Mathilde Buisson[4], Guillaume Christophe[2], Giuseppe D'Auria[5,6], Florence De Taroni[2], Vincent Foulongne[7,8], Rémy Froissart[1], Christelle Graf[3], Sophie Grasset[1,2], Soraya Groc[1,7], Christophe Hirtz[9], Audrey Jaussent[10], Julie Lajoie[11], Frédérique Lorcy[12], Eric Picot[2], Marie-Christine Picot[10], Jacques Ravel[13], Jacques Reynes[11], Thérèse Rousset[12], Aziza Seddiki[4], Martine Teirlinck[2], Vincent Tribout[2], Édouard Tuaillon[8], Tim Waterboer[15] , Nathalie Jacobs[16], Ignacio G Bravo[1], Michel Segondy[7,8], Nathalie Boulle[8,12] and Samuel Alizon[1,*]

For peer review only

Word count: 5265 words, excluding title page, abstract, references, figures and tables.

1 Laboratoire MIVEGEC (UMR 5290 CNRS, IRD, UM), 911, avenue Agropolis, BP 64501, 34394 Montpellier, France
2 Center for Free Information, Screening and Diagnosis (CeGIDD), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
3 Department of Obstetrics and Gynaecology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France.
4 Department of Research and Innovation (DRI), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
5 Sequencing and Bioinformatics Service, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública), 46020 Valencia, Spain
6 CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain
7 Department of Virology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
8 Pathogenesis and Control of Chronic Infections, INSERM, CHU, University of Montpellier, Montpellier, France
9 University of Montpellier, LBPC/PPC- IRMB, CHU de Montpellier, 80 rue Augustin Fliche, Montpellier, France
10 Department of Medical Information (DIM), Centre Hospitalier Universitaire de Montpellier, Montpellier, France
11 Department of Medical microbiology, University of Manitoba, 745 Bannatyne, Winnipeg, Canada
12 Department of pathology and oncobiology, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
13 Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, USA
14 Department of Infectious and Tropical Diseases, Centre Hospitalier Universitaire de Montpellier, Montpellier, France
15 German Cancer Research Center (DKFZ), Infections and Cancer Epidemiology, Im Neuenheimer Feld 280, Heidelberg, Germany
16 GIGA-Research, Cellular and molecular immunology, University of Liège, 3 Avenue de l'Hôpital, 4000 Liège, Belgium

* Author for correspondence: samuel.alizon@cnrs.fr

For peer review only

Abstract

Introduction

Human papillomaviruses (HPVs) are responsible for one third of all cancers caused by infections. Most HPV studies focus on chronic infections and cancers, and we know little about the early stages of the infection. Our main objective is to better understand the course and natural history of cervical HPV infections in healthy, unvaccinated and vaccinated, young women, by characterising the dynamics of various infection-related populations (virus, epithelial cells, vaginal microbiota, and immune effectors). Another objective is to analyse HPV diversity within hosts and in the study population, in relation to co-factors (lifestyle characteristics, vaccination status, vaginal microbiota, human genetics).

Methods and Analysis

The PAPCLEAR study is a mono-centric longitudinal study following 150 women, aged 18-25 years, for up to 2 years. Visits occur every 2 or 4 months (depending on HPV status) during which several variables are measured, such as behaviours (via questionnaires), vaginal pH, HPV presence and viral load (via qPCR), local concentrations of cytokines (via MesoScale Discovery technology) and immune cells (via flow cytometry). Additional analyses are outsourced, such as titration of circulating anti-HPV antibodies, vaginal microbiota sequencing (16S and ITS-1 loci) and human genotyping. To increase the statistical power of the epidemiological arm of the study, an additional 150 women are screened cross-sectionally. Finally, to maximise the resolution of the time series, participants are asked to perform weekly self-samples at home. Statistical analyses will involve classical tools in epidemiology, genomics, and virus kinetics, and will be performed or coordinated by the CNRS in Montpellier.

Ethics and Dissemination

This study has been approved by the CPP Sud Méditerranée I (reference number 2016-A00712-49); by the CCTIRS (reference number 16.504); by the CNIL (reference number MMS/ABD/AR1612278, decision number DR-2016-488) and by the ANSM (reference 20160072000007). Results will be published in preprint servers, peer-reviewed journals and disseminated through conferences.

Trial registration number: NCT02946346

For peer review only

Keywords: HPV; acute infection; persistence; virus load; immunity; microbiota; viral kinetics

Article summary

Strengths and limitations of this study

- Short time interval between the visits (every two months for infected women) and additional self-sampling every week at home.

- The combination of virological (virus load), immunological (cytokine concentrations and immune cell percentages) and environmental (vaginal microbiota composition, pH) measurements at each visit.

- A limitation is that the density of the follow-up limits the number of participants (N=150), which can affect the power of epidemiological analyses.

- We complement the longitudinal study with a cross-sectional study of N=150 women to increase statistical power.

For peer review only

Introduction

Epidemiology of HPV genital infections in young adults and public health implications

Infections by Human Papillomaviruses (HPVs) are likely the most common sexually transmitted infection (STI) globally. It is often estimated that, worldwide, more than 80% of sexually-active individuals will be infected by an HPV type [1]. In France, a study performed in 2013 in the Paris area estimated the prevalence of HPV genital infections to be 25% in women below 25 years of age [2]. In the area of Montpellier (France), the prevalence of oncogenic HPVs (also referred to as 'high-risk', HR, HPVs) in pregnant women aged 16 to 42 years was close to 20% [3]. These numbers are consistent with worldwide estimates according to which HPVs are most prevalent in women under 25 years of age, with an estimated overall prevalence of 24% [4].

Fortunately, the vast majority of infections by HPVs are asymptomatic and benign. Even for HPV16, which is probably the most oncogenic human virus, only a minority of infections (less than 10%) become persistent [5], and then a minority of these (12%) progress to cancer if untreated [1, 6]. Indeed, it is estimated that approximately 70 to 100% of infections by HPVs are cleared within 12 to 24 months, with strong differences between virus types [5, 7–9]. Recent studies suggest that primo-infections could be shorter in young girls [10] but, in general, there are many unknowns about the biology of non-persisting infections [11].

Our lack of knowledge partly comes from the fact that in vaccine trials, from which most of the data on infection duration originate, participants are followed every six months for several years [5, 7, 9, 12]. This frequency is sufficient to estimate the time to clearance (or persistence) but it is not precise enough to understand the within-host dynamics, often referred to as 'kinetics' [13], of infections that last on average 6 to 24 months. Arbitrarily, after 24 months, an infection is often considered to be persistent [14].

Some factors have been shown to correlate with persistence (e.g. immunosuppression, smoking, and co-infection with other STIs [15]) but we do not know how these affect viral kinetics. Also, some changes in viral-immunity interactions appear to be related to persistence

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

and disease progression [16–19] but, again, we do not know the underlying interactions between the viruses, the host target cells, and the immune response in acute infections [11]. Finally, it has been argued that the vaginal microbiota may differ between HPV-infected and HPV-uninfected women [20] and that specific microbiota composition may interact with HPV detection [21]. However, it is difficult to disentangle the cause and the consequence. For instance, does the microbiota composition change after the establishment of an HPV infection, or do certain microbiota compositions increase susceptibility to HPV infection?

A better understanding of the within-host infection dynamics and of the determinants of clearance and persistence of viral infection is particularly important in the context of vaccination [22–25]. Indeed, the long-term efficacy of the anti-HPVs vaccines at the population level will largely depend on the within-host viral dynamics because, ultimately, most selective pressures on viral populations occur via the immune response [26]. Furthermore, a better understanding of acute HPV infections can shed new light on issues related to latency, fertility, or immunotherapies [11].

Prevention strategies and treatment

Treatment

Since most infections by HPVs are benign in young adults and clear within six to 24 months, the current standard of care is to avoid over-treatment, even in the presence of cervical lesions [27]. Clinical interventions (colposcopies, biopsies, and surgery) are less often performed with young women (< 25 years) and only for high-grade (pre-cancerous) lesions (cervical intraepithelial neoplasia grade 2, CIN-2, or more). Low-grade lesions (CIN-1) are not systematically treated but rather monitored yearly to detect any progression to high-grade lesions.

Genital warts caused by non-oncogenic HPVs (often referred to as ‘low-risk’, LR, HPVs) can be removed by surgery or treated with bi- and trichloroacetic acid, cryotherapy or other treatments [28].

For peer review only

HPV vaccination

There are currently three licensed vaccines: a bivalent vaccine (Cervarix®) targeting HPV16 and HPV18 (together accounting for 70% of cervical cancers [1]), a quadrivalent vaccine (Gardasil®) that additionally targets HPV6 and HPV11 (non-oncogenic, but highly prevalent and associated to benign proliferative lesions) and, since 2014, a nonavalent vaccine (Gardasil 9®) that targets five more oncogenic types (HPV31, HPV33, HPV45, HPV52, and HPV58, which altogether account for 20% of cervical cancers [24]). These vaccines succeed in eliciting a protective immune response against new infections by the targeted viruses, and are used throughout the world, albeit with wide variation in coverage (for reviews, see e.g. [29, 30]).

Vaccination campaigns in France started in 2006 but with limited coverage: it reached 28.5% in 2008 [31] and has been decreasing ever since [32]. The vaccine is recommended for girls from 11 to 14 years of age, currently with a vaccination scheme of two doses with a six months interval. A catch-up is organised for girls aged 15-19 years, with a three-doses vaccination scheme. Vaccination is reimbursed by the French Social Security but is not mandatory. It is also recommended for men who have sex with men (MSM) as well as for immuno-compromised people [32]. Vaccination is now the primary prevention strategy against cervical cancers.

Screening

In France, the secondary prevention strategy against cervical cancer is routine individual cytology-based screening for pre-cancerous and cancerous cervical lesions in women between 25 and 65 years old. Cytology can also be performed in younger women if they report risk factors for cervical cancer (multiple partners, chronic STIs or HIV infection [32]). Detection of oncogenic HPVs is proposed for triage in case of abnormal cytology (i.e. high-grade or low-grade squamous intraepithelial lesion, HSIL and LSIL respectively, or Atypical Squamous Cells of Undetermined Significance, ASCUS).

Primary objectives

The first primary objective of this cohort study is to decipher the kinetics and ecology of cervical HPV infections in healthy young women, i.e. follow the population dynamics of the virus, the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

target epithelial cells, and the immune effectors.

The second primary objective is to characterise the diversity of genital HPVs in young women in the region of Montpellier in relationship with their lifestyle, vaccination status, vaginal microbiota, and human genetics.

Secondary objectives

A secondary objective is to characterise the acquisition and clearance dynamics of cervical HPV infections as a function of viral diversity, host immunity, vaginal microbiota and human genetics.

A final objective is to investigate variations in genetic diversity of HPVs during cervical infections.

Methods and analysis

Participants

The study population is composed of young women at risk of HPV infection. The age class was chosen because it exhibits high HPV prevalence (24% worldwide [4] and approximately 25% in France [2]). Inclusion of younger women would have raised technical issues because of the requirement for parental consent.

Women are recruited through a social media page, and through posters and leaflets distributed at the Universities in Montpellier and at the Montpellier STI screening centre (*Centre Gratuit d'Information de Dépistage et de Diagnostic*, CeGIDD). The composition of the population visiting the CeGIDD has already been documented in an earlier study [33]. In total, the centre is visited by approximately 3,000 women per year, the majority of which are under 25 years of age (80%). Approximately 40% of the attendants report three or more partners over the last twelve months and approximately 50% report using adequate behaviour for prevention against HIV.

For peer review only

Since all participants are healthy, they are referred to as participants rather than patients. As in any longitudinal study, ensuring participant commitment is challenging. To achieve this goal, we have set up a compensation of 40 EUR per visit and an additional 10 EUR in case of a complete follow-up. Furthermore, participants who have answered a sufficient number of questionnaires and brought back a sufficient number of self-samples get a 100 EUR bonus at the end. Overall, a participant performing 12 return visits would gain a total compensation of 650 EUR.

Inclusion criteria

Participants are women from 18 to 25 years old living in the metropolitan area of Montpellier. They must be sexually active with at least one new partner over the last 12 months. This criteria is fixed to maximise the incidence of new HPV infections. Participants must be able to and willing to give written informed consent: they must sign an informed consent form, understand the requirements for the study, and be affiliated to a French social security scheme (which is a state requirement).

Women cannot be included in the study if they have a history of HPV-associated pathology (genital warts or cervical lesions), if they are pregnant or intending to become pregnant in the coming year, infected by HIV, undergoing (or planning to undergo) intense medical treatment (biotherapy, chemotherapy, immunosuppression), planning on moving outside the Montpellier metropolitan area within the next 18 months, in a dependency or employment with the sponsor or the investigator, if they participated in a clinical trial involving administration of drugs within the last four weeks or if they belong to a vulnerable group (e.g. children, adults with physical or mental disabilities).

For peer review only

Design/setting

This study has a longitudinal component aimed at deciphering within-host dynamics and a cross-sectional component aimed at understanding the diversity of HPV infections in young adults in the area of Montpellier, France. The general structure of the study is shown in Fig 1.

If a woman fits the main inclusion criteria, she can go through an inclusion visit ($V1$) with a physician (gynaecologist or midwife) at the CeGIDD. During this visit, the study investigator presents the study and checks all inclusion criteria before asking the participant to read and sign the informed consent form. Participants then undergo a medical consultation during which a number of samples are collected (see below). They then fill out health and lifestyle questionnaires and are given cotton-flocked swabs for self-sampling at home the next visit, along with instructions on how to fill in weekly questionnaires through an online form (these are performed throughout the study).

An appointment is scheduled four weeks later for the Results visit ($V2$), where the cervical cytology results are communicated. Additional samples are collected and self-sample swabs for home collection are provided.

The next return visits (Vi, where $i > 2$) are as follows:

- Participants with a positive DEIA HPV test (see below), i.e. infected by an *Alphapapillomavirus*, at $V1$ join the HPV positive (HPV+) arm of the study with return visits scheduled every 2 months.
- Participants with a negative DEIA HPV test at $V1$ join the HPV negative (HPV-) arm with return visits scheduled every 4 months.
- HPV- participants infected by an *Alphapapillomavirus* move to the HPV+ arm.

Intervals between visits are based on earlier results showing that HPV infections last from 9 to 18 months on average depending on the HPV type [5, 7–9] and that a total follow-up of 4 months yields results that are difficult to analyse [21]. The longer interval in the HPV- arm is based on the estimated incidence for HPV genital infections in young women, which is greater

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

than 30% [34, 35].

Participants in the HPV- arm are followed until month 32 of the study.

Participants in the HPV+ arm are followed until they clear the infection or until they have been infected for 24 months (after which we consider that the infection is persistent). Clearance is defined as being negative at two visits in a row for the first HPV type detected in the follow-up.

In between these visits to the CeGIDD, participants are asked to perform regular (every week for HPV+ and every second week for HPV-) self-samples using vaginal swabs, to measure vaginal pH and to fill in a short questionnaire. Self-samples are stored in the participants' freezer and brought back at every visit.

The study will end with the last HPV+ participant having cleared the infection or been infected for 24 months.

Patients and public involvement

Participants did not play a role in the design of this study and the

results of the study will be disseminated to participants who have left the study and to the general public via an email newsletter in French.

Visits

The summary of the visit schedule and of the samples collected at each visit is shown in Table 1.

Inclusion visit (V_1)

This visit takes place at the CeGIDD and is scheduled by the Clinical Research Technician

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(CRT) via phone or email.

Women meet a study investigator, who explains the goals and requirements of the study and checks that the inclusion criteria are met. If so, after a general discussion, the informed consent forms are signed.

A female physician/midwife performs a general exam and then a gynaecological exam during which the following samples are taken:

- vaginal pH cotton swab (EcoCare™),
- vaginal swab (Copan ESwab™) in 1mL Amies liquid for DNA extraction and microbiota analysis,
- vaginal swab (Copan ESwab™) in 1mL of RNA preservation medium,
- ophthalmic sponge (Weck-cel®) to collect cervical secretions for cytokines analysis,
- cervical smear in 20mL of Thinprep® (Preservcyt® liquid) for HPV and HSV assays, and cytology evaluation.

Following the gynaecological consultation, the participant meets with a nurse to measure body temperature, blood pressure and draw 20mL of blood (a 5mL tube for SNPs sequencing, a 10mL tube for immunophenotyping and a 5mL tube for HPV antibody titration). For the longitudinal study, the nurse provides the participant with 3 self-sampling kits, 3 pH strips, a freezer box to bring back to the next visit, as well as instructions on how to perform the home sampling and store the samples in her personal freezer until the next visit.

If the participant has not been tested for a STI in the last 3 months, the nurse draws an additional blood tube of 5mL to test for STIs (HIV, HCV, HBV) and collects vaginal self-samples for chlamydiae and gonorrhea detection. Syphilis testing is prescribed to participants who meet the STI clinic's guidelines.

Finally, the participant meets with the CRT to fill in questionnaires #1 (inclusion visit) and #3 (home). The CRT answers any remaining questions, explains how to fill the home

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

questionnaires (#3) and sets an appointment for the Results visit.

Results visit (V_2)

During this visit, the participants are given the result of cervical lesion screening using the liquid cytology (normal, ASCUS, LSIL or HSIL). Participants with a high-grade lesion (HSIL) exit the study and are referred to the gynaecology service of the CHU of Montpellier.

During this visit, the physician/midwife collects additional samples: 2 vaginal swabs for DNA and RNA analysis, and a cervical smear in 10mL of PBS (to confirm HPV status and perform flow cytometry analyses).

The participant fills in questionnaires #2 (for return visits) and #3 (home). An appointment for the next visit is set and swabs for home self-sampling are given.

Return visits (V_i)

These visits only occur in the longitudinal study.

HPV- arm. Participants uninfected by HPV visit the clinic every 4 months until month 26. During these visits, the same samples as in the inclusion visit (V_1**)** are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

The nurse only draws blood if a screening test for STIs other than HPV is required. The participant then fills in questionnaires #2 and #3 and an appointment is set for the next visit in 16 weeks.

If an HPV infection is detected in the cervical smear collected during this visit, the participant moves to the HPV+ arm and the CRT contacts the participant to move her appointment forward.

HPV+ arm. Participants infected by HPV visit the clinic every 2 months. They cannot switch arm and will remain in the HPV+ arm until clearance or the end of the study. During the visits, the same samples as in the inclusion visit (V_0) are collected by the physician/midwife except for the cervical smear, which is put in PBS instead of Thinprep.

For peer review only

The nurse then draws 5mL of blood for HPV antibody titration. If this is the first HPV+ visit following an HPV- visit, the nurse also draws 10mL of blood for immunophenotyping. Finally, if a test for additional STIs is needed, the nurse draws 5mL of blood and collects vaginal self-samples for STI detection.

Importantly, if the participant has been infected by a HR-HPV for more than 12 months and cytology has not been performed within the last 12 months, the cervical smear is put in Thinprep® fixation medium, instead of PBS, for cytological analysis (cervical lesion screening).

Finally, the participant fills in questionnaires #2 and #3, receives self-samples for home collection and an appointment is set for the next visit in 8 weeks.

Endpoints

The primary endpoint for the study is the inclusion and follow-up of HPV-infected women in order to describe the kinetics of HPV virus load, and the associated immune response.

Secondary endpoints are the characterisation of the interactions between the course of the infection (e.g. duration), the HPV type(s), the abundance and taxonomic diversity of bacteria, fungi and viruses in the vaginal microbiota, human genetics (SNPs) and basal immunological status.

Technical procedures

DNA extraction

DNA extraction from cervical smears will be performed using Nuclisens EasyMAg from Biomerieux or an equivalent protocol. For the microbiota analyses, special kits involving physical (via beads) and/or enzymatic breaking of the cellular barrier will be favoured following standard protocols to study the vaginal microbiome [36], e.g. the MagAttract®

For peer review only

PowerMicrobiome® DNA/RNA kit from Qiagen.

HPV detection, typing and quantification

The participants' infection status (HPV+ or HPV-) will be assessed using the DEIA test, which is based on a PCR of the short SPF10 amplicon [37] and detects all *Alphapapillomaviruses* with great sensitivity.

If the DEIA test is positive, HPVs will be typed using the LiPA25 kit, which is based on the same SPF10-PCR, and has a lower detection threshold compared to other hybridisation-based typing methods [38].

The reason for basing the detection on the DEIA rather than the LiPA25 is that some *Alphapapillomavirus* may be detected by DEIA but not genotyped by LiPA and also that the DEIA is more sensitive than the LiPA. If the DEIA is positive and the LiPA25 is negative, typing will be performed by sequencing the product of a PGMY09/11 PCR [39], which targets another region of the HPV genome than the SPF10 PCR.

The quantification of HPV DNA genome copy number in the samples will be performed using the protocol set up by Micalessi et al. [40].

Cytokine titration

Cytokines can be used as markers of immune activation or immunosuppression and can also inform us on which components of the immune system are involved. Cervical sponges are centrifuged after the addition of 175μL of PBS. Cervical secretions are analysed for a set of 5 to 6 cytokines levels using the Meso Scale Discovery (MSD) Multiplex ELISA platform, which has a low detection threshold and a slowly saturating dose-response curve. Based on earlier results [41, 42], we will first investigate a large panel of 20 cyctokines (IFN-α2a, IFN-γ, IL-1α, IL-5, IL-6, IL-8, IL-10, IL-12, IL-15, IL-17, IL-18, IL-23, IL-25, IP-10, MCP-1, MIP-1α, MIP-3α, MIP-3β, TNF-α, TNF-β) to choose the most relevant ones for a longitudinal follow-up.

Flow cytometry

Analysing immune cells via flow cytometry is extremely challenging on cells as fragile as the

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ones from cervical smears. However, several studies suggest that this is feasible [41–43]. Here, we follow the protocol described in [44].

Staining is performed using a Duraclone custom mix targeting CD45, CD3, CD4, CD8, CD16, CD56, CD69, CD161 and TCR$\gamma\delta$. The last marker, Live&Dead tests for cellular viability. Samples are acquired using a Navios flow cytometer (Beckman Coulter, three-laser configuration).

Sequencing

Sequencing will be performed for microbiota profiling. It involves PCR amplification of the V3-V4 region of 16S RNA for bacteria [45] and ITS1 for fungi [46]. We anticipate that the bacteria should belong to the operational taxonomic units (OTU) described in the five community state types found in vaginal communities [47, 48]. The virome will also be explored using shotgun sequencing and rolling circle PCR amplification [49]. Human genetics are explored using chip sequencing for SNPs.

Statistical analyses

Times series analyses

The core results of the study will come from the longitudinal follow-up of infected women, which will generate time series, i.e. a set of values collected from the same individual over time (Figure 2). There will be several time series per individual (virus load, number of immune cells, cytokine and antibody levels). These time series will be used to fit mathematical viral kinetics models that describe the interaction between viruses, host target cells (here, in the case of HPV, keratinocytes) and the immune response. These models are commonly developed for viral infections [13, 50–52], including those caused by HPVs [53]. We anticipate our follow-up to yield adequate data for such a fit based on the estimated duration of HPV infections (9 to 18 months [5, 7–9]). Furthermore, the weekly self-samples allow us to increase the resolution if necessary.

We will use non-linear mixed effect models [54] to jointly analyse time series from all

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

participants. More precisely, we will rely on *R* packages such as nlme [55] or lme4 [55]. Note that, in addition to estimating model parameters (e.g. life-expectancy of infected cells or virion production rate of infected cells), this approach also allows us to compare biological models using statistical tools based on model likelihood such as Akaike Information Criterion. For an example of such analysis in the case of HIV, see [51].

Microbiota dynamics

The composition of the vaginal microbiota has already been described and shown to exhibit considerably less diversity than the gut microbiota [47]. The dynamics of this microbiota has also been studied and shown to closely follow menstrual cycles [48].

Using the time series of OTU abundances (measured via 16S RNA sequencing and qPCR) we will infer interaction parameters by assuming an underlying Lotka-Volterra competition model [56]. This work will include time series analysis techniques (e.g. auto-correlation or local similarity analysis) and statistical inference methods in order to infer community structure and interactions from the next-generation sequencing (NGS) datasets [57]. Finally, statistical methods from ecology will also be used to study community diversity (e.g diversity indices) and community assembly, such as cluster and ordination analyses [58].

Genome Wide Association Studies

We will use human single nucleotide polymorphisms (SNPs) inferred by chip sequencing to look for genetic determinants of key traits (e.g. microbiota composition or HPV infection duration). This is classically done by performing a Genome Wide Association Study (GWAS), which is a complex regression method designed for situations where there are many explanatory variables (here millions of SNPs) for a single trait of interest. GWAS will be performed using classical methods [59]. Earlier GWAS studies have been applied to HPV infections for instance to test for determinants to the ability to seroconvert following infection [60] and cervical cancer (see [61] for a review). Here, our expected sample (N = 300 women) is limited but SNPs with large effects have been detected by studies with comparable sizes [62].

For peer review only

Additional analyses

For all collected variables, descriptive statistics will be calculated according to the level of measurement. For metric variables these statistics can be mean and standard deviation as well as quantiles and more robust statistics [63]. In case of categorical variables group proportions and contingency tables are prepared.

Univariate inferential statistics follow a descriptive analysis. Generally, parametric testing procedures are preferred to non-parametric tests, as the former have higher power. That is why, for metric variables, we will first check whether the data can be assumed to be normally distributed. For normally distributed variables, ANOVA statistics are done to detect differences between groups. In case of significance, post-hoc analysis (Tukey test) are planned to reveal pairwise differences. If the data are not normally distributed or ordinally scaled, non-parametric analyses will be used. These contain the Kruskal-Wallis test and the Wilcoxon test as a post-hoc test with an appropriate correction of the significance level. Since the cell counts are expected to be small, Fisher's exact test will be performed for contingency tables instead of the asymptotic χ^2 test for categorical variables.

Sample size calculation

The study will enrol a total of $N = 300$ women, with $N = 150$ in a longitudinal study and $N = 150$ in a cross-sectional study. The goal of the longitudinal study is to follow 75 women longitudinally, preferentially before they are infected (see above). For the following calculations, we assumed a high percentage of lost during follow-up (30%).

With 150 enrolments and considering that the prevalence of HPV infection in young women is ≈ 60% (based on our preliminary data) and 30% of lost to follow-up, we expect to detect (and successfully follow) 63 infections at inclusion [CI95: 51−75, assuming a binomial distribution to calculate the 95% confidence interval].

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Among women who are uninfected at the first visit and considering the yearly incidence being close to 30% [64], we expect 12 [CI95: 6−20] to be infected during the first year of follow-up (still assuming 30% of lost to follow-up).

In the end, with 150 enrolments and assuming a high percentage of lost to follow-up (30%), we expect to successfully follow 75 [CI95: 56−95] women infected at different stages of HPV infection: beginning, during and end.

This will be made possible by the probability of transmission of HPV, which is estimated to be ≈ 90% without condom use and still high with condom use (≈ 40%) [34].

Finally, regarding potential interference with the HPV vaccines, we do not anticipate any significant problem for two reasons. First, as mentioned above, the vaccine coverage is low in France [32]. Second, and more importantly, the vaccines only target few HPV types, thus leaving open the possibility of infection by dozens of types. Furthermore, studying the kinetics of a non-vaccine HPV type in a vaccinated woman will be extremely informative, e.g. to detect any potential cross-reactivity [65].

To run cross-sectional analyses (especially on the microbiota and human genetics), we will enrol *N* = 150 additional women who will only perform the inclusion and the results visits. This sample size was chosen to reach that of earlier GWAS studies [61, 62].

Trial governance

Sponsor

This study is sponsored by the Centre Hospitalier Universitaire (CHU) of Montpellier. The CHU is involved in the implementation of the trial, legal/ethical submissions (see below for details on Ethics approval) and implementing the clinical database (eCRF), which is hosted by Ennov-Clinical (ClinSight). The CHU is not involved in the analysis or interpretation of the data. The CHU of Montpellier performs regular quality control assessments. A clinical research assistant

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

will visit the CeGIDD every 4 months to ensure that implementation is in accordance with the protocol. The CHU has taken out insurance from the Société hospitalière d'assurances mutuelles, 18, rue Edouard Rochet-6 9372 Lyon cedex 08 (contract number 138983) through the full research period, covering its own civil liability and that of any agent (clinical or research staff), in accordance with article L.1121-10 of the French Public Health Code.

Scientific committee

The scientific committee comprises the study investigators, clinicians, scientific experts and representatives of the sponsor. The committee meets yearly and is responsible for following research progress, monitoring compliance with good clinical practices and patient safety. It can also decide relevant modification of the protocol. Requests from third parties to access data collected during the study will be evaluated by the committee.

Monitoring

Monitoring is performed during the whole study at CeGIDD according to the sponsor specific SOP. Routine monitoring visits are made by the monitors designated by the sponsor to check compliance with the protocol, the completeness, accuracy and consistency of the data, and adherence to GCP. The principal investigator ensures that eCRFs are completed in a timely manner and must allow periodical access to eCRFs, patient records, drug logs, and all other study-related documents and materials. The frequency of monitoring visits is determined by factors such as study design and the site enrolment requirements but visits will normally occur at least once every 4 months.

Trial registration

The trial has been registered to ClinicalTrials.gov on 27 Oct 2016 with ID number NCT02946346.

For peer review only

Ethics and Dissemination

The PAPCLEAR trial obtained favourable opinions from the Comité de Protection des Personnes (CPP) Sud Méditerranée I on May 11, 2016 (CPP number 16 42, reference number ID RCB 2016-A00712-49); from the Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (CCTIRS) on July 12, 2016 (reference number 16.504); and from the Commission Nationale Informatique et Libertés (CNIL) on Dec 16, 2016 (reference number MMS/ABD/AR1612278, decision number DR-2016-488). This trial was authorised by the Agence Nationale de Sécurité du Médicament et des Produits de Santé (ANSM) on July 20, 2016 (reference 20160072000007).

The protocol has been modified since its initial version and the latest modification was accepted by the CPP on Dec 12, 2018.

All participants in the study will sign an informed consent form prior to participation.

The results will be published on preprint servers (e.g. BioRxiv), peer-reviewed journals, post-print servers (e.g. HAL) and disseminated through conferences.

Discussion

Expected results

Acute infections by HPVs are important to study because vaccination is most effective when performed before the first infection. However, we currently know very little about the early stages of HPV infections. This clinical study will give us an unprecedented level of detail into the natural history of HPV infections in young women. Variations in virus load over time have been studied but in the context of cervical cancer in older women [66]. In addition, we will also describe the nature and the dynamics of the immune response (local immune cells and cytokines, circulating anti-HPV antibodies) and of the vaginal microbiota. Beyond these kinetics, we will also have access to data such as infection clearance or not in 24 months, presence of

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

more than one HPV type or coinfection by other STIs.

These data will be analysed in the light of numerous cofactors. One of the most important will be human genetics, with the sequencing of millions of SNPs. Others will be related to the sexual behaviour (number of partners, contraception methods, sexual practices) and general lifestyle. We, therefore, expect broader insights regarding sexual health in young women.

Practical and operational issues

One of the main practical challenges resides in the analysis of cervical smears by flow cytometry. Indeed, the tissues are known to be fragile, adhesive and auto-fluorescent. Even though standard protocols now exist [44], they require processing fresh samples in less than 2 hours.

Another potential issue has to do with contamination by HPV DNA between samples, which are frequent in the HPV field due to the robustness of the virions and the sensitivity of the tests. To certify our ability to control for these, we have entered the 2017 GLOBAL HPV DNA Proficiency Panel from the WHO HPV LabNet [67].

Regarding the enrolment of the participants, we do not expect issues with enrolling 150 women in 28 months for the longitudinal study and 150 for the cross-sectional study. This is due to the number of visitors of the centre who fit the inclusion criteria (more than 3,000 per year) and because of earlier high participation rates in the same population ([33] enrolled 1381 participants in 5 months for their study).

Concerning the follow-up, the high incidence rate of HPV can also lead to transient carriage, i.e. women who are positive for a type only at a single visit. This has been observed for instance in longitudinal studies with a tight follow-up interval [21]. To control for this, we will run the HPV detection test on the cells from the cervical smear after washing with RPMI.

For peer review only

Table 1: Summary of the visit schedules and samples take. The cross-sectional study only includes the first two columns (V1 and V2). The ¤indicate samples taken at visits. + participants infected by a HR-HPV for 12 month will have one PBS smear replaced by a Thinprep® smear to perform a cytology and check for lesions. ⊲ this sample is only taken at the first HPV+ visit of a formerly HPV- participant. ⋆ STI detection will be performed at inclusion unless the participant has been tested within the last 3 months and during the study every 6 months if a new partner has been reported or upon request.

For peer review only

	Inclusion ($V1$)	Results ($V2$)	Return (Vi, with $i > 2$)	
Participants	all	all	HPV+	HPV-
Time	day 0	+ 4 weeks	+ 8 weeks	+ 16 weeks
Eligibility	¤			
Consent	¤			
Gynecological consult	¤	¤	¤	¤
Vaginal pH coton swab	¤	¤	¤	¤
2 vaginal swab samples (Copan ESwabTM)	¤	¤	¤	¤
1 ophtalmological sponge sample	¤		¤	¤
1 cervical smear in Thinprep® (cytology)	¤		+	
1 cervical smear in PBS		¤	+	¤
Blood sampling (HPV antibodies)	¤		¤	
Blood sampling (sequencing)	¤			
Blood sampling (immunophenotyping)	¤		◅	
Other STI detection	⋆	⋆	⋆	⋆
Questionnaire #1 (inclusion)	¤			
Questionnaire #2 (visit)		¤	¤	¤
Questionnaire #3 (home)	¤	¤	¤	¤
Returning self-sampling samples		¤	¤	¤
Serious Adverse Event collection		¤	¤	¤

Abbreviations

STI: sexually transmitted infection

TCR: T-cell receptor,

WHO: World Health Organisation. CNRS: Centre National de la Recherche Scientifique,

CRT: Clinical Research Technician,

ELISA: enzyme-linked immunosorbent assay,

GWAS: Genome Wide Association Study,

HIV: Human Immunodeficiency Virus,

HPV: Human Papillomavirus,

HR: high-risk,

ITS: Internal Transcribed Spacer,

HSIL: High grade Squamous Intraepithelial Lesion,

LR: low-risk,

LSIL: Low grade Squamous Intraepithelial Lesion,

NGS: Next Generation Sequencing,

OTU: Operational Taxonomic Unit,

PBMC: Peripheric Blood Mononuclear Cell,

PBS: Phosphate Buffered Saline,

RPMI: Roswell Park Memorial Institute medium,

SNP: Single Nucleotide Polymorphism, ANOVA: Analysis of variance,

ASC-US: Atypical squamous cells of undetermined significance,

CD: Cluster of differentiation,

CI95: 95% Confidence interval,

CeGIDD: Centre Gratuit d'Information de Dépistage et de Diagnostic,

CHU: Centre Hospitalier Universitaire,

CIN: Cervical intraepithelial Neoplasia,

For peer review only

Trial status

The study began on Oct 1, 2016 and the first inclusion was on Nov 3, 2016. On Jun 23, 2018, 89 participants have been included in the longitudinal study. Inclusions in the longitudinal study will continue until March 2019 and the study is expected to last until Aug 2021.

Conflicts of interests

The authors have read and understood BMJ policy on declaration of interests and declare that they have no competing interests.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 648963). The authors acknowledge further support from the Centre National de la Recherche Scientifique (CNRS), the Institute de Recherche pour le Développement (IRD) and the Centre Hospitalier Universitaire (CHU) of Montpellier.

Data statement

All personal and identifying information collected from participants are kept in a secure place at the CeGIDD during the duration of the trial and will be destroyed at the end of the study. The final raw dataset will be accessible only to the sponsor (CHU) and the chief scientist's (SA) team. Anonymous data will be available to external parties upon approval of both the sponsor and the scientific committee. All publications will be made green or gold open access and the corresponding data will be provided as supplementary material or via a public repository (e.g. Zenodo), provided that there is no conflict with ethical guidelines.

Author contributions

Samuel Alizon, Carmen Lia Murall and Massical Rahmoun were the major contributors in the conception of the protocol. Samuel Alizon wrote the initial version of the manuscript. Christian

For peer review only

Selinger, Monique Baldellou, Claire Bernat, Marine Bonneau, Vanina Boué, Mathilde Buisson, Guillaume Christophe, Giuseppe D'Auria, Florence De Taroni, Vincent Foulongne, Rémy Froissart, Christelle Graf, Sophie Grasset, Soraya Groc, Christophe Hirtz, Audrey Jaussent, Julie Lajoie, Frédérique Lorcy, Eric Picot, Marie-Christine Picot, Jacques Ravel, Jacques Reynes, Thérèse Rousset, Aziza Seddiki, Martine Teirlinck, Vincent Tribout, Édouard Tuaillon, Tim Waterboer, Nathalie Jacobs, Ignacio G Bravo, Michel Segondy and Natalie Boulle were involved in the conception of the protocol, in the implementation of the study and read and approved the final manuscript.

For peer review only

Acknowledgements

We thank all the study participants and the CeGIDD staff for their commitment to the study. We also thank the reviewers and, in particular, Dr. Andrew Brouwer for his meticulous reading of the manuscript.

References

1. Tota, J.E., Chevarie-Davis, M., Richardson, L.A., Devries, M., Franco, E.L.: Epidemiology and burden of HPV infection and related diseases: implications for prevention strategies. Prev Med **53 Suppl 1**, 12–21 (2011). doi:10.1016/j.ypmed.2011.08.017

2. Monsonego, J., Zerat, L., Syrjänen, K., Zerat, J.C., Smith, J.S., Halfon, P.: Prevalence of genotype-specific HPV infection among women in France: implications for screening and vaccination. Gynecol Obstet Fertil **41**(5), 305–313 (2013). doi:10.1016/j.gyobfe.2013.03.003

3. Brun-Micaleff, E., Coffy, A., Rey, V., Didelot, M.-N., Combecal, J., Doutre, S., Daurès, J.-P., Segondy, M., Boulle, N.: Cervical cancer screening by cytology and human papillomavirus testing during pregnancy in french women with poor adhesion to regular cervical screening. J Med Virol **86**(3), 536–45 (2014). doi:10.1002/jmv.23764

4. Bruni, L., Diaz, M., Castellsagué, X., Ferrer, E., Bosch, F.X., de Sanjosé, S.: Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. J Infect Dis **202**(12), 1789–99 (2010). doi:10.1086/657321

5. Insinga, R.P., Dasbach, E.J., Elbasha, E.H., Liaw, K.-L., Barr, E.: Incidence and duration of cervical human papillomavirus 6, 11, 16, and 18 infections in young women: an evaluation from multiple analytic perspectives. Cancer Epidemiol Biomarkers Prev **16**(4), 709–15 (2007). doi:10.1158/1055-9965.EPI-06-0846

6. Woodman, C.B.J., Collins, S.I., Young, L.S.: The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer **7**(1), 11–22 (2007). doi:10.1038/nrc2050

7. Rodríguez, A.C., Schiffman, M., Herrero, R., Wacholder, S., Hildesheim, A., Castle, P.E., Solomon, D., Burk, R., Proyecto Epidemiológico Guanacaste Group: Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections. J Natl Cancer Inst **100**(7), 513–7 (2008). doi:10.1093/jnci/djn044

For peer review only

8. Trottier, H., Mahmud, S., Prado, J.C.M., Sobrinho, J.S., Costa, M.C., Rohan, T.E., Villa, L.L., Franco, E.L.: Type-Specific Duration of Human Papillomavirus Infection: Implications for Human Papillomavirus Screening and Vaccination. J Infect Dis **197**(10), 1436–1447 (2008). doi:10.1086/587698

9. Ramanakumar, A.V., Naud, P., Roteli-Martins, C.M., de Carvalho, N.S., de Borba, P.C., Teixeira, J.C., Blatter, M., Moscicki, A.-B., Harper, D.M., Romanowski, B., Tyring, S.K., Ramjattan, B., Schuind, A., Dubin, G., Franco, E.L.: Incidence and duration of type-specific human papillomavirus infection in high-risk HPV-naïve women: results from the control arm of a phase II HPV-16/18 vaccine trial. BMJ Open **6**(8), 011371 (2016). doi:10.1136/bmjopen-2016-011371

10. Houlihan, C.F., Baisley, K., Bravo, I.G., Kapiga, S., de Sanjosé, S., Changalucha, J., Ross, D.A., Hayes, R.J., Watson-Jones, D.: Rapid acquisition of HPV around the time of sexual debut in adolescent girls in Tanzania. Int J Epidemiol **45**(3), 762–773 (2016). doi:10.1093/ije/dyv367

11. Alizon, S., Murall, C.L., Bravo, I.G.: Why Human Papillomavirus Acute Infections Matter. Viruses **9**(10), 293 (2017). doi:10.3390/v9100293

12. Herrero, R., Wacholder, S., Rodríguez, A.C., Solomon, D., González, P., Kreimer, A.R., Porras, C., Schussler, J., Jiménez, S., Sherman, M.E., Quint, W., Schiller, J.T., Lowy, D.R., Schiffman, M., Hildesheim, A., Costa Rica Vaccine Trial Group: Prevention of persistent human papillomavirus infection by an HPV16/18 vaccine: a community-based randomized clinical trial in Guanacaste, Costa Rica. Cancer Discov **1**(5), 408–19 (2011). doi:10.1158/2159-8290.CD-11-0131

13. Canini, L., Perelson, A.S.: Viral kinetic modeling: state of the art. J Pharmacokinet Pharmacodyn **41**(5), 431–443 (2014). doi:10.1007/s10928-014-9363-3

14. Stanley, M.: Immune responses to human papillomavirus. Vaccine **24**(S1), 16–22 (2006).

For peer review only

doi:10.1016/j.vaccine.2005.09.002

15. Ferenczy, A., Franco, E.: Persistent human papillomavirus infection and cervical neoplasia. Lancet Oncol **3**(1), 11–6 (2002)

16. zur Hausen, H.: Review: Papillomaviruses — to Vaccination and Beyond. Biochemistry **73**(5), 498–503 (2008). doi:10.1134/S0006297908050027

17. Einstein, M.H., Schiller, J.T., Viscidi, R.P., Strickler, H.D., Coursaget, P., Tan, T., Halsey, N., Jenkins, D.: Clinician's guide to human papillomavirus immunology: knowns and unknowns. Lancet Infect Dis **9**(6), 347–56 (2009). doi:10.1016/S1473-3099(09)70108-2

18. Van Hede, D., Langers, I., Delvenne, P., Jacobs, N.: Origin and immunoescape of uterine cervical cancer. Presse Med **43**(12P2), 413–421 (2014). doi:10.1016/j.lpm.2014.09.005

19. Stanley, M.: Immunology of HPV infection. Curr Obstet Gynecol Rep **4**(4), 195–200 (2015). doi:10.1007/s13669-015-0134-y. Accessed 2017-03-20

20. Gao, W., Weng, J., Gao, Y., Chen, X.: Comparison of the vaginal microbiota diversity of women with and without human papillomavirus infection: a cross-sectional study. BMC Infect Dis **13**, 271 (2013). doi:10.1186/1471-2334-13-271

21. Brotman, R.M., Shardell, M.D., Gajer, P., Tracy, J.K., Zenilman, J.M., Ravel, J., Gravitt, P.E.: Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. J Infect Dis **210**(11), 1723–33 (2014). doi:10.1093/infdis/jiu330

22. Koutsky, L.A., Ault, K.A., Wheeler, C.M., Brown, D.R., Barr, E., Alvarez, F.B., Chiacchierini, L.M., Jansen, K.U., Proof of Principle Study Investigators: A controlled trial of a human papillomavirus type 16 vaccine. N Engl J Med **347**(21), 1645–51 (2002). doi:10.1056/NEJMoa020586

23. Riethmuller, D., Jacquard, A.-C., Lacau St Guily, J., Aubin, F., Carcopino, X., Pradat, P., Dahlab, A., Prétet, J.-L.: Potential impact of a nonavalent hpv vaccine on the occurrence of hpv-

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

related diseases in france. BMC Public Health **15**, 453 (2015). doi:10.1186/s12889-015-1779-1

24. Joura, E.A., Giuliano, A.R., Iversen, O.-E., Bouchard, C., Mao, C., Mehlsen, J., Moreira, E.D. Jr, Ngan, Y., Petersen, L.K., Lazcano-Ponce, E., Pitisuttithum, P., Restrepo, J.A., Stuart, G., Woelber, L., Yang, Y.C., Cuzick, J., Garland, S.M., Huh, W., Kjaer, S.K., Bautista, O.M., Chan, I.S.F., Chen, J., Gesser, R., Moeller, E., Ritter, M., Vuocolo, S., Luxembourg, A., Broad Spectrum HPV Vaccine Study: A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. N Engl J Med **372**(8), 711–23 (2015). doi:10.1056/NEJMoa1405044

25. Murall, C.L., Bauch, C.T., Day, T.: Could the human papillomavirus vaccines drive virulence evolution? Proc Biol Sci **282**, 20141069 (2015). doi:10.1098/rspb.2014.1069

26. Alizon, S., Méthot, P.-O.: Reconciling Pasteur and Darwin to control infectious diseases. PLoS Biol **16**(1), 2003815 (2018). doi:10.1371/journal.pbio.2003815

27. Moscicki, A.-B., Ma, Y., Wibbelsman, C., Darragh, T.M., Powers, A., Farhat, S., Shiboski, S.: Rate of and Risks for Regression of CIN-2 in adolescents and young women. Obstet Gynecol **116**(6), 1373–1380 (2010). doi:10.1097/AOG.0b013e3181fe777f

28. Buck Jr., H.W.: Warts (genital). BMJ Clin Evid **2015**, 1602 (2015)

29. Herrero, R., González, P., Markowitz, L.E.: Present status of human papillomavirus vaccine development and implementation. Lancet Oncol **16**(5), 206–16 (2015). doi:10.1016/S1470-2045(14)70481-4

30. Maver, P.J., Poljak, M.: Progress in prophylactic human papillomavirus (HPV) vaccination in 2016: A literature review. Vaccine (2018). doi:10.1016/j.vaccine.2017.07.113

31. Fagot, J.-P., Boutrelle, A., Ricordeau, P., Weill, A., Allemand, H.: HPV vaccination in France: uptake, costs and issues for the National Health Insurance. Vaccine **29**(19), 3610–6 (2011). doi:10.1016/j.vaccine.2011.02.064

For peer review only

32. Ben Hadj Yahia, M.-B., Dervaux, B., Duport, N., Floret, D., Gaillot, J., Heard, I., Jacquet, A., Le Goaster, C., Lévy-Bruhl, D., Morer, I., Parent du Chatelet, I., Peigue-Lafeuille, H., Rumeau-Pichon, C.: Vaccination contre les infections à papilloamvirus. Technical report, Haut Conseil de la Santé Publique, Paris, France (2014). https://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=454

33. Clarivet, B., Picot, E., Marchandin, H., Tribout, V., Rachedi, N., Schwartzentruber, E., Ledésert, B., Dereure, O., Guillot, B., Picot, M.-C.: Prevalence of Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium in asymptomatic patients under 30 years of age screened in a French sexually transmitted infections clinic. Eur J Dermatol **24**(5), 611–6 (2014). doi:10.1684/ejd.2014.2413

34. Winer, R.L., Hughes, J.P., Feng, Q., O'Reilly, S., Kiviat, N.B., Holmes, K.K., Koutsky, L.A.: Condom use and the risk of genital human papillomavirus infection in young women. N Engl J Med **354**(25), 2645–54 (2006). doi:10.1056/NEJMoa053284

35. Winer, R.L., Hughes, J.P., Feng, Q., Stern, J.E., Xi, L.F., Koutsky, L.A.: Incident Detection of High-Risk Human Papillomavirus Infections in a Cohort of High-Risk Women Aged 25-65 Years. J Infect Dis **214**(5), 665–75 (2016). doi:10.1093/infdis/jiw074

36. Ravel, J., Brotman, R.M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D.W., Sakamoto, J., Koenig, S.S., Fu, L., Zhou, X., Hickey, R.J., Schwebke, J.R., Forney, L.J.: Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. Microbiome **1**(1), 29 (2013). doi:10.1186/2049-2618-1-29

37. Kleter, B., van Doorn, L.-J., ter Schegget, J., Schrauwen, L., van Krimpen, K., Burger, M.,

For peer review only

ter Harmsel, B., Quint, W.: Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses. Am J Pathol **153**(6), 1731–1739 (1998). doi:10.1016/S0002-9440(10)65688-X

38. Geraets, D.T., Struijk, L., Kleter, B., Molijn, A., van Doorn, L.-J., Quint, W.G.V., Colau, B.: The original SPF10 LiPA25 algorithm is more sensitive and suitable for epidemiologic HPV research than the SPF10 INNO-LiPA Extra. J Virol Meth **215-216**, 22–29 (2015). doi:10.1016/j.jviromet.2015.01.001

39. Gravitt, P.E., Peyton, C.L., Alessi, T.Q., Wheeler, C.M., Coutlée, F., Hildesheim, A., Schiffman, M.H., Scott, D.R., Apple, R.J.: Improved amplification of genital Human Papillomaviruses. J Clin Microbiol **38**(1), 357–361 (2000)

40. Micalessi, I.M., Boulet, G.A.V., Bogers, J.J., Benoy, I.H., Depuydt, C.E.: High-throughput detection, genotyping and quantification of the human papillomavirus using real-time PCR. Clin Chem Lab Med **50**(4), 655–61 (2012). doi:10.1515/cclm.2011.835

41. Hunter, P.J., Sheikh, S., David, A.L., Peebles, D.M., Klein, N.: Cervical leukocytes and spontaneous preterm birth. Journal of Reproductive Immunology **113**, 42–49 (2016). doi:10.1016/j.jri.2015.11.002

42. Shannon, B., Yi, T.J., Perusini, S., Gajer, P., Ma, B., Humphrys, M.S., Thomas-Pavanel, J., Chieza, L., Janakiram, P., Saunders, M., Tharao, W., Huibner, S., Shahabi, K., Ravel, J., Rebbapragada, A., Kaul, R.: Association of HPV infection and clearance with cervicovaginal immunology and the vaginal microbiota. Mucosal Immunology **10**(5), 1310–1319 (2017). doi:10.1038/mi.2016.129

43. Lajoie, J., Juno, J., Burgener, A., Rahman, S., Mogk, K., Wachihi, C., Mwanjewe, J., Plummer, F.A., Kimani, J., Ball, T.B., Fowke, K.R.: A distinct cytokine and chemokine profile at the genital mucosa is associated with HIV-1 protection among HIV-exposed seronegative commercial sex workers. Mucosal Immunol **5**(3), 277–287 (2012). doi:10.1038/mi.2012.7

For peer review only

44. Juno, J.A., Boily-Larouche, G., Lajoie, J., Fowke, K.R.: Collection, isolation, and flow cytometric analysis of human endocervical samples. J Vis Exp **89**, 51906 (2014). doi:10.3791/51906

45. Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J.: Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol **74**(8), 2461 (2008). doi:10.1128/AEM.02272-07

46. Findley, K., Oh, J., Yang, J., Conlan, S., Deming, C., Meyer, J.A., Schoenfeld, D., Nomicos, E., Park, M., NIH Intramural Sequencing Center Comparative Sequencing Program, Kong, H.H., Segre, J.A.: Topographic diversity of fungal and bacterial communities in human skin. Nature **498**(7454), 367–370 (2013). doi:10.1038/nature12171. Accessed 2017-09-13

47. Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., Forney, L.J.: Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA **108**, 4680–7 (2011). doi:10.1073/pnas.1002611107

48. Gajer, P., Brotman, R.M., Bai, G., Sakamoto, J., Schütte, U.M.E., Zhong, X., Koenig, S.S.K., Fu, L., Ma, Z.S., Zhou, X., Abdo, Z., Forney, L.J., Ravel, J.: Temporal dynamics of the human vaginal microbiota. Sci Transl Med **4**(132), 132–52 (2012). doi:10.1126/scitranslmed.3003605

49. Johne, R., Müller, H., Rector, A., van Ranst, M., Stevens, H.: Rolling-circle amplification of viral DNA genomes using phi29 polymerase. Trends Microbiol **17**(5), 205–211 (2009). doi:10.1016/j.tim.2009.02.004

50. Nowak, M.A., May, R.M.: Virus Dynamics: Mathematical Principles of Immunology and Virology. Oxford University Press, Oxford, USA (2000)

51. Stafford, M.A., Corey, L., Cao, Y., Daar, E.S., Ho, D.D., Perelson, A.S.: Modeling plasma virus concentration during primary HIV infection. J. theor. Biol. **203**(3), 285–301 (2000).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

doi:10.1006/jtbi.2000.1076

52. Perelson, A.S.: Modelling viral and immune system dynamics. Nat. Rev. Immunol. **2**(1), 28–36 (2002). doi:10.1038/nri700

53. Murall, C.L., Jackson, R., Zehbe, I., Boulle, N., Segondy, M., Alizon, S.: Epithelial stratification shapes infection dynamics. PLoS Comput Biol **15**(1), 1006646 (2019). doi:10.1371/journal.pcbi.1006646

54. Steimer, J.L., Vozeh, S., Racine Poon, A., Holford, N., O'Neil, R.: The population approach: rationale, methods and applications in clinical pharmacology and drug development. In: Balant, P.G.W..L. (ed.) Handbook of Experimental Pharmacology, vol. 110, pp. 405–451. Springer, Berlin (1994)

55. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. Journal of Statistical Software **67**(1) (2015). doi:10.18637/jss.v067.i01

56. Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., Deng, L., Yeliseyev, V., Delaney, M.L., Liu, Q., Olle, B., Stein, R.R., Honda, K., Bry, L., Gerber, G.K.: MDSINE: Microbial dynamical systems INference engine for microbiome time-series analyses. Genome Biology **17**, 121 (2016). doi:10.1186/s13059-016-0980-6. Accessed 2017-03-09

57. Faust, K., Lahti, L., Gonze, D., de Vos, W.M., Raes, J.: Metagenomics meets time series analysis: unraveling microbial community dynamics. Curr Opin Microbiol **25**, 56–66 (2015). doi:10.1016/j.mib.2015.04.004

58. Fox, G.A., Negrete-Yankelevich, S., Sosa, V.J.: Ecological Statistics: Contemporary Theory and Application. Oxford University Press, Oxford, USA (2015)

59. Shi, Y., Li, L., Hu, Z., Li, S., Wang, S., Liu, J., Wu, C., He, L., Zhou, J., Li, Z., Hu, T., Chen, Y., Jia, Y., Wang, S., Wu, L., Cheng, X., Yang, Z., Yang, R., Li, X., Huang, K., Zhang, Q., Zhou, H., Tang, F., Chen, Z., Shen, J., Jiang, J., Ding, H., Xing, H., Zhang, S., Qu, P., Song, X., Lin,

For peer review only

Z., Deng, D., Xi, L., Lv, W., Han, X., Tao, G., Yan, L., Han, Z., Li, Z., Miao, X., Pan, S., Shen, Y., Wang, H., Liu, D., Gong, E., Li, Z., Zhou, L., Luan, X., Wang, C., Song, Q., Wu, S., Xu, H., Shen, J., Qiang, F., Ma, G., Liu, L., Chen, X., Liu, J., Wu, J., Shen, Y., Wen, Y., Chu, M., Yu, J., Hu, X., Fan, Y., He, H., Jiang, Y., Lei, Z., Liu, C., Chen, J., Zhang, Y., Yi, C., Chen, S., Li, W., Wang, D., Wang, Z., Di, W., Shen, K., Lin, D., Shen, H., Feng, Y., Xie, X., Ma, D.: A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. Nat Genet **45**(8), 918–22 (2013). doi:10.1038/ng.2687

60. Chen, D., Gaborieau, V., Zhao, Y., Chabrier, A., Wang, H., Waterboer, T., Zaridze, D., Lissowska, J., Rudnai, P., Fabianova, E., Bencko, V., Janout, V., Foretova, L., Mates, I.N., Szeszenia-Dabrowska, N., Boffetta, P., Pawlita, M., Lathrop, M., Gyllensten, U., Brennan, P., McKay, J.D.: A systematic investigation of the contribution of genetic variation within the MHC region to HPV seropositivity. Hum Mol Genet **24**(9), 2681–2688 (2015). doi:10.1093/hmg/ddv015

61. Chen, D., Gyllensten, U.: Lessons and implications from association studies and post-GWAS analyses of cervical cancer. Trends Genet **31**(1), 41–54 (2015). doi:10.1016/j.tig.2014.10.005

62. Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., Cozzi-Lepri, A., De Luca, A., Easterbrook, P., Francioli, P., Mallal, S., Martinez-Picado, J., Miro, J.M., Obel, N., Smith, J.P., Wyniger, J., Descombes, P., Antonarakis, S.E., Letvin, N.L., McMichael, A.J., Haynes, B.F., Telenti, A., Goldstein, D.B.: A whole-genome association study of major determinants for host control of HIV-1. Science **317**(5840), 944–947 (2007). doi:10.1126/science.1143767

63. Huber, P.J.: The 1972 Wald Lecture Robust Statistics: A Review. Ann Math Stat **43**(4), 1041–1067 (1972). doi:10.1214/aoms/1177692459

64. Winer, R.L., Feng, Q., Hughes, J.P., O'Reilly, S., Kiviat, N.B., Koutsky, L.A.: Risk of female human papillomavirus acquisition associated with first male sex partner. J Infect Dis

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

197(2), 279–82 (2008). doi:10.1086/524875

65. Herrero, R.: Human Papillomavirus (HPV) Vaccines: Limited Cross-Protection against Additional HPV Types. J Infect Dis **199**(7), 919–922 (2009). doi:10.1086/597308

66. Depuydt, C.E., Verstraete, L., Berth, M., Beert, J., Bogers, J.-P., Salembier, G., Vereecken, A.J., Bosmans, E.: Human papillomavirus positivity in women undergoing intrauterine insemination has a negative effect on pregnancy rates. Gynecol Obstet Invest **81**(1), 41–6 (2016). doi:10.1159/000434749

67. WHO HPV LabNet. World Health Organization. http://www.who.int/biologicals/areas/human_papillomavirus/WHO_HPV_LabNet/en/

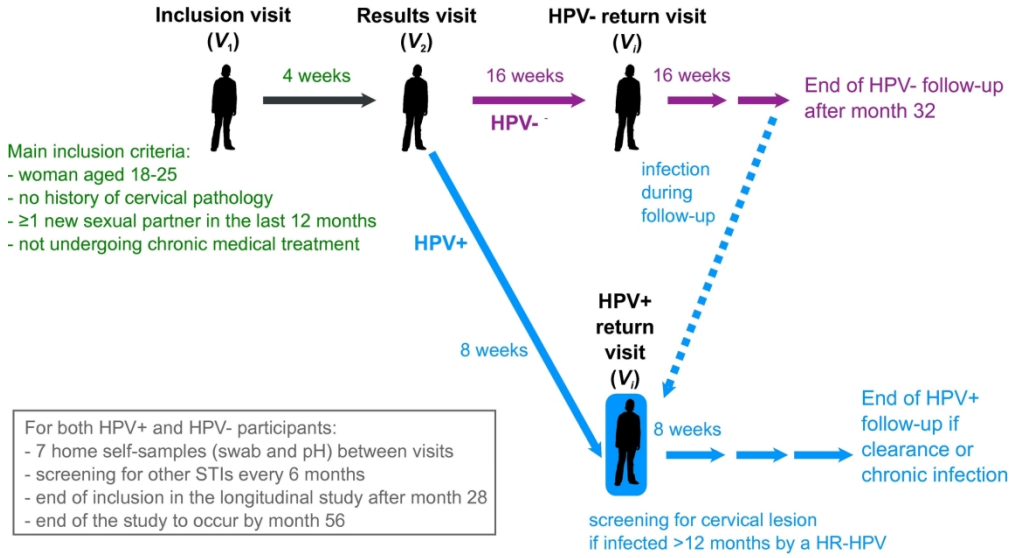
For peer review only

Figure captions

Figure 1: General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V_1), a results visit (V_2) and then return visits (V_i with $i > 2$). For the cross-sectional study, participants only have V_1 and V_2.

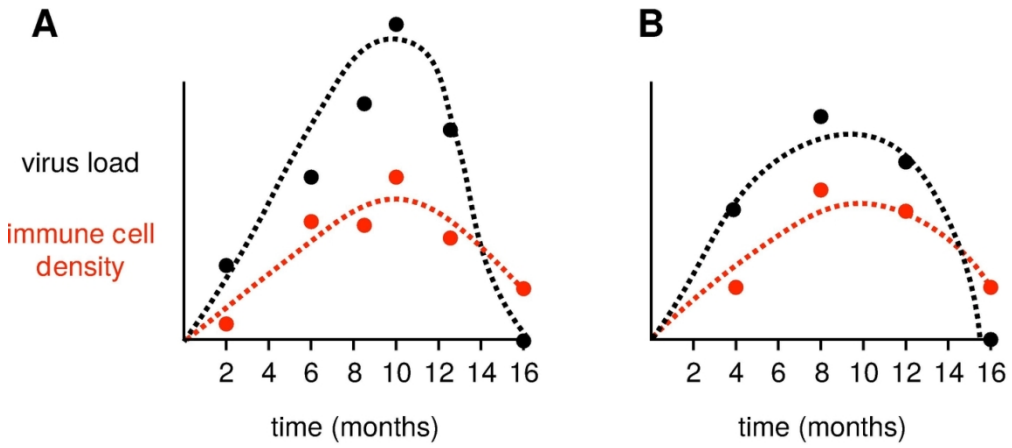
Figure 2: Fitting viral kinetics models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

For peer review only



General structure of the PAPCLEAR study. For the longitudinal study, participants have an inclusion visit (V1), a results visit (V2) and then return visits (Vi with i > 2). For the cross-sectional study, participants only have V1 and V2.

190x104mm (300 x 300 DPI)



Fitting kinetics dynamical models to within-host times series. Dashed lines indicate a model fitted using virus load (in black) or immune cells (in red) time series. In panel A, the follow-up is bi-monthly with 2 missing visits and several delayed visits, whereas in panel B the follow-up is every 4 months without any missing or delayed visits. In spite of missing data this, the situation shown in panel A is clearly the best for inferring parameter values and for fitting the underlying dynamics.

120x52mm (300 x 300 DPI)