BMJ Open

Laser capture microdissection as a tool to evaluate human papillomavirus genotyping and methylation as biomarkers of persistence and progression of anal lesions

Alyssa M Cornall,1,2 Jennifer M Roberts,3 Monica Molano,1 Dorothy A Machalek,1,2 Samuel Phillips,1,2 Richard J Hillman,4 Andrew E Grulich,5 Fengyi Jin,5 I Mary Poynten,5 David J Templeton,5,6 Suzanne M Garland,1,7 Sepehr N Tabrizi,1,7 on behalf of the SPANC Study Team

ABSTRACT

Introduction: Anal squamous cell carcinoma is preceded by persistent infection with high-risk human papillomavirus (HPV) and the cancer precursor, high-grade squamous intraepithelial lesion (HSIL). Detection of specific HPV genotypes and HPV-related biomarkers may be an option for primary anal screening. However, more data on the natural history of HPV-related anal lesions are required. The outcomes from this study will enhance our understanding of the clinical and biological behaviour of HPV-related anal lesions and inform the development of future HPV genotype and/or biomarker screening tests.

Methods and analysis: HIV-negative and HIV-positive men who have sex with men, aged 35 years and over, recruited from community-based settings in Sydney, Australia, attend 6 clinic visits over 3 years. At the first 5 visits, participants undergo a digital anorectal examination, an anal swab for HPV genotyping and anal cytology, and high-resolution anoscopy with directed biopsy of any visible abnormalities that are suggestive of any abnormality suspicious of SIL. Tissue sections from participants diagnosed with histologically confirmed HSIL at the baseline clinic visit will undergo laser capture microdissection, HPV detection and genotyping, and quantitation of CpG methylation in baseline and follow-up biopsies. Histological and cytological findings in combination with HPV genotyping data will be used to identify persistent HSIL. HSIL will be stratified as non-persistent and persistent based on their status at 12 months. The performance of HPV genotype and methylation status in predicting disease persistence at 12 months will be assessed, along with associations with HIV status and other covariates such as age.

Ethics and dissemination: The St Vincent's Hospital Ethics Committee granted ethics approval for the study. Written informed consent is obtained from all individuals before any study-specific procedures are performed. Findings from this study will be disseminated to participants and the community through study newsletters, and through peer-reviewed publications and international conferences.

Strengths and limitations of this study

▪ One of only a very small number of cohorts with prospectively collected longitudinal anal cytology and histology samples from HIV-positive and HIV-negative participants.
▪ Participants are not routinely treated for high grade squamous intraepithelial lesion, allowing for longitudinal follow-up of lesions.
▪ Use of laser capture microdissection to isolate lesion tissue.
▪ Challenges relating to sampling and diagnostic uncertainty and establishing whether or not a lesion is persistent from one visit to the next.
▪ Some lesions will be missed during high resolution anoscopy; therefore, the study uses both histological diagnoses and cytological findings.

INTRODUCTION

Human papillomavirus (HPV)-associated anal squamous cell carcinoma (ASCC), like cervical cancer, is preceded by persistent infection with high-risk HPV (HR-HPV) and precancerous high-grade squamous intraepithelial lesions (HSILs). In developed nations, rates of ASCC in both genders have been increasing for decades.1 Subpopulations at greatest risk of ASCC include men who have sex with men (MSM), HIV-positive men and women, solid organ transplant recipients on immunosuppressant drugs, and women with previous HPV-related cervical, vulvar or vaginal disease.2 MSM have a 20-fold
increased incidence. HIV-positive people have at least a 30-fold increased incidence, and anal cancer has become the most commonly occurring non-AIDS-defining cancer in developed nations. HIV-positive MSM are the group at highest risk, with 50-fold to 150-fold increased incidence.\(^5\)\(^6\) Women with previous HPV-related anogenital disease have a 5-fold to 20-fold increased risk of anal cancer.\(^7\) and solid organ transplant recipients have a 5-fold increased incidence.\(^8\)

Vaccination against HPV infection is likely to reduce anal cancer rates in the long term, but not within the next few decades. The quadrivalent HPV (qHPV) vaccine (targeting HPV 6, 11, 16 and 18) has been shown in phase III clinical trials to greatly reduce the risk of development of anal cancer precursor lesions in MSM.\(^9\)\(^10\) The introduction of prophylactic qHPV vaccination of young Australian females in 2007 led to a 77% reduction in prevalence of vaccine preventable HPV types,\(^11\) as well as a reduction in the incidence of cervical precancerous lesions in young women.\(^12\)\(^13\) Vaccination of Australian high school-aged boys with the qHPV vaccine began in 2013.\(^14\) A new nonavalent HPV vaccine, targeting an extended range of genotypes (HPV 6, 11, 16, 18, 31, 33, 45, 52 and 58) has recently completed clinical trials.\(^15\) This vaccine has been approved for use by the US Food and Drug Administration (FDA),\(^16\) and decisions about the use of this vaccine are yet to be made for general and at-risk populations, and ongoing studies are generating data to inform these decisions.

Currently available HPV vaccines are not therapeutic and do not protect against disease development in individuals already HPV DNA-positive for vaccine genotypes,\(^17\)\(^18\) although individuals who are seropositive and DNA-negative are protected from future disease.\(^19\)

Because vaccination does not protect against existing ongoing infections, it may not be as effective in adults, and consequently widespread vaccination of adults over the age of 26 years has not been adopted by any country to date. To reduce ASCC morbidity in the unvaccinated populations, other approaches are necessary. Organised population-based cytology screening programmes have led to a continuing decline in incidence and mortality from cervical cancer in Australia and other industrialised nations, where high standards of quality assurance and quality control are in place and there is high coverage of the target population.\(^20\) Epidemiological, aetiological and pathological similarities between squamous cancers of the cervix and anus have led to suggestions that similar screening programmes in high-risk groups could also reduce the incidence of ASCC.\(^21\)\(^22\) However, it has been established that anal cytology has not been shown to be as effective as cervical cytology in terms of screening to prevent cancer.\(^21\) It has also been determined that genotyping of anal swabs is not predictive for lesion-specific genotypes.\(^23\)

Existing data indicate that approximately 20–40% of MSM have anal HSIL at any one time,\(^4\) with only a very small proportion progressing to cancer even in the absence of treatment.\(^24\) Therefore, the specificity of anal cytological screening is likely to be low. More specific screening targets, such as DNA or RNA from the most high-risk genotypes, and/or other biomarkers of cancer risk, may be a necessary part of any screening programme.

Between 70% and 90% of ASCC are associated with HPV type 16 in single or mixed infections; the contributions of other HPV genotypes are not well established, but may predominantly be restricted to a small subset of causative genotypes.\(^25\)\(^26\) In line with ongoing changes to cervical screening practices in Australia and the USA,\(^27\)\(^28\) detection of specific HPV genotypes and HPV-related biomarkers, such as methylation, HPV mRNA transcript patterns and protein markers, may be options for primary anal screening. Host-specific biomarkers such as increased methylation in regulatory regions of tumour suppressor genes have also shown promise as a potential triage marker for cervical high-grade disease,\(^29\) and there are some data to support this finding in the anal canal.\(^30\)

HPV DNA-based anal screening algorithms and vaccination policy for anal cancer prevention in high-risk populations will rely on more accurate HPV genotype data than are currently available. Genotype data for anal cancers are limited; however, they suggest that the distribution of HPV genotypes associated with anal cancers is different to that observed in cervical cancers, with a higher proportion of HPV16-positive ASCC and a higher, but still low, proportion of low-risk HPV (LR-HPV) and possible/probable HR-HPV genotypes.\(^31\)\(^32\) Reliable genotype data for anal HSIL are even more limited than for cancers. In addition to the limited quantity of data,\(^33\)\(^36\) the interpretation of these data are confounded by the common occurrence of multiple HPV genotypes and multiple squamous lesions in the anal canals of MSM\(^37\) and, to a lesser extent, women.\(^38\)

While each individual lesion or cancer is caused by a single genotype,\(^25\)\(^39\) up to a quarter of anal cancer specimens from men and women\(^25\)\(^39\) and up to half of anal HSIL biopsies contain multiple detectable genotypes,\(^25\) most of which will be incidental to the lesion of interest. Consequently, attributions of genotype to lesions, based on swab or biopsy samples, are indirect and often rely on the use of mathematical algorithms to determine the most likely causative genotype.\(^39\)\(^42\)

A more direct method, laser capture microdissection (LCM) combined with sensitive genotype testing, is a highly effective approach to assign HPV genotypes to individual lesions.\(^25\)\(^39\) LCM utilises lasers to precisely isolate and capture specifically targeted tissue, from a single cell to a selected segment of an intraepithelial lesion, from which nucleic acid can be extracted and tested. We describe here the use of LCM to study the natural history of longitudinally collected HPV-associated anal lesions. The data generated using the methods described below will enable us to accurately determine which HPV genotypes present the highest risk of anal
HSIL persistence and/or progression to cancer, investigate patterns of HPV-related biomarkers related to persistence of HSIL, and inform the development of screening and vaccine strategies for the prevention of anal cancer in at-risk populations.

METHODS AND ANALYSIS

Participant recruitment
Approximately 600 MSM, 35 years and older, recruited from community-based settings in Sydney, Australia, are being enrolled into this 5-year cohort study. Participant recruitment and sample collection have been previously described in detail. Once recruited, participants attend a baseline clinic visit and four follow-up visits at 6, 12, 24 and 36 months after baseline. LCM is performed on lesions collected at baseline and at follow-up visits from consenting participants with biopsy-confirmed HSIL diagnosed at baseline.

Sampling of anal lesions
At each clinic visit, participants undergo a digital anorectal examination, intra-anal swabbing for HPV DNA genotyping as well as anal cytology, which is followed by high-resolution anoscopy (HRA) and biopsy of any lesions suspicious of SIL as described elsewhere. Clinical impressions of intra-anal abnormalities and their anatomical positions observed during HRA are recorded on a preformatted form by the physician and digital photographs of the area are taken. All clinical reporting by the anoscopist is carried out with the knowledge of the participants’ HIV status, cytology and HRA results from previous visits. Biopsies are placed into 10% formalin solution and sent to Douglass Hanly Moir Pathology (Sydney, Australia) where they are routinely processed and embedded in paraffin. Formalin-fixed paraffin-embedded (FFPE) biopsies are sectioned and stained with H&E, and an initial diagnosis obtained, as described. Reporting of the biopsies is performed in accordance with criteria, terminology and recommendations of the Lower Anogenital Squamous Terminology (LAST) Project. In addition to the LAST criteria, anal low-grade squamous intraepithelial lesions (LSILs) are routinely subdivided into exophytic (eLSIL) and flat (fLSIL), based on architectural criteria (eLSIL corresponding to condyloma acuminatum).

HPV genotyping of anal cytology samples
A 1 ml aliquot of intra-anal swab material in ThinPrep PreservCyt (Hologic, Inc, Marlborough, Massachusetts, USA) is centrifuged at 13,000×g for 20 min and the cell pellet resuspended in 200 µL of sterile phosphate-buffered saline. DNA is extracted using the MagNApure LC (Roche Diagnostics GmbH, Penzberg, Germany) using the DNA-I kit (blood cells high-performance protocol) or on the MagNApure 96 using the DNA and Viral Nucleic Acid Small Volume Kit (Pathogen Universal 200 protocol). Extracted DNA is eluted into a final volume of 100 µL. HPV genotyping is performed on all samples using the Linear Array HPV Genotyping Test (Roche Diagnostics), which detects and identifies 37 of the most common mucosal HPV genotypes. Samples which return an unassessable result on Linear Array are retested using DNA diluted 1 in 2 in nuclease-free water, to dilute any substances which may be inhibiting PCR amplification. Samples which test positive for HPV52/33/35/58 and one or more of HPV33, 35 and/or 58 are further tested on a quantitative PCR (qPCR) assay to determine whether HPV52 is also present, as previously described. Samples that are negative for HPV16 or 18 will be tested on a type-specific qPCR assay for these genotypes, as previously described.

Selection of biopsies
Participants diagnosed with biopsy-confirmed HSIL at the baseline visit and a further 50 participants diagnosed with LSIL are included for comparison. Participants are excluded if no tissue remains in the biopsy block, the biopsy is negative for any SIL on review, or if the participant did not give consent for their tissue to be stored and used for procedures other than histological diagnosis. Participants who have had ablative treatment are also excluded. More than one biopsy may be selected from a single participant.

Follow-up of HSILs
By reference to clinical notes and photographs taken during HRA (figure 1A), biopsies from follow-up visits where available are selected for sectioning. A HSIL is considered persistent if it is (1) detected at least 12 months after the baseline visit within the same octant or up to one octant either side of the baseline lesion (figure 1B); (2) where possible confirmed by reference to photographs taken during HRA and annotated by a clinician; and (3) if it contains the same HPV genotype following LCM and genotyping. An anal intraepithelial neoplasia grade 2 (AIN2) could be considered the same lesion as an AIN3 diagnosed at a subsequent visit, and vice versa, as both are HSIL. An LSIL diagnosed on follow-up is not necessarily considered to be the same lesion as a HSIL diagnosed at a previous visit; however, exceptions may include low-grade lesions that are p16-positive and positive for the same HPV genotype, if an eligible HSIL is detected at the next visit. In cases where it is not clear whether a HSIL is still present at a follow-up visit, additional data including cytology results and reference to HRA photographs will be used in accordance with the algorithm in figure 2. HSIL will be defined as persistent for <6 or <12 (not persistent), or ≥12 months (persistent).

Sandwich sectioning and diagnosis of FFPE samples for LCM
For each biopsy, a series of nine sections are cut and processed for histological analysis using a sandwich sectioning method modified from refs 39 and 49 (figure 3). To minimise cross-contamination, the
microtome blade is changed and the microtome surface is cleaned with xylene after sectioning of each sample. Where possible without resetting the blocks, biopsies are cut from the basal side towards the epithelial/superficial side to minimise carryover of unrelated viral DNA to lesional tissue. The two outer sections (sections 1 and 9) are cut to 3 µm thickness and stained with H&E. The first inner section (section 2) is also cut to 3 µm thickness and stained with antibodies specific for the presence of the virus.

Figure 1 (A) Photograph of the anal transformation zone taken through a speculum during high-resolution anoscopy. A digital photograph at 10x magnification of an area treated with 3% acetic acid. An area suspicious for high-grade disease is outlined; this was confirmed by biopsy to be anal intraepithelial neoplasia grade 3 (AIN3). (B) Clock face diagram of anal transformation zone segmented into eight quadrants. A high-grade squamous intraepithelial lesion is considered to be potentially persistent if it was biopsied from within the same quadrant or up to one quadrant either side of the baseline lesion at follow-up clinic visits.

Figure 2 Flow chart for defining probable persistence of high-grade squamous intraepithelial lesion (HSIL) detected at baseline. A HSIL biopsy detected at consecutive visits which tests positive for the same human papillomavirus (HPV) genotype following laser capture microdissection (LCM) and is located in the same octant or an octant either side of the baseline biopsy is considered persistent. In the event that not all of these criteria are met, additional information including the cytology result and swab HPV genotype result and extent of disease on high-resolution anoscopy (HRA) will be considered. A HSIL lesion is considered to be not persistent if there were no follow-up biopsy taken at 6 months meeting the criteria for persistence, and if the cytology and swab results also do not support persistence. If there is doubt, 12-month samples will be taken into consideration and if not found to support the existence of a persistent HSIL, the result will be 'not persistent'.
70% ethanol. The tissue is air dried to ensure that no residual ethanol is present and is subsequently incubated with 400 µg of proteinase K in 200 µL of tissue lysis buffer (Roche Molecular Systems) on a 55°C heat block for 1 h followed by 37°C overnight. Samples which are not fully digested are vortexed briefly and incubated for a further hour following the addition of a further 10 µL of proteinase K (10 mg/mL in tissue lysis buffer). The digested tissue is then transferred onto an automated MagnaPure 96 (Roche) isolation and purification system, using the DNA and Viral NA small volume kit, Universal Pathogen 200 protocol. The final volume of eluted DNA in solution is 100 µL.

**Dewaxing, LCM and DNA extraction of lesions**

Immediately prior to tissue capture by LCM, PEN membrane slides are de-paraffinised via sequential 5 min incubations in 100% xylene (twice) and 100% ethanol (twice) and then allowed to dry thoroughly prior to laser capture. An annotated photograph of the H&E-stained section is used to guide selection of abnormal tissue to be captured. Tissue sections are visualised and dissected with an ultraviolet cutting laser on a Veritas 704 Laser Capture Microdissection System (Arcturus Bioscience, Mountain View, California, USA) and a single lesion is captured onto each thermoplastic CapSure Macro LCM Cap (Applied Biosystems) with an infrared capture laser (figure 4). Care is taken to avoid capturing the superficial layers of the epithelium (figure 5). For any lesions which test positive for more than one HPV genotype (genotype testing described below), additional sections will be captured by LCM from the basal layer at either end of the originally dissected area of tissue to test for multiple colliding lesions or contamination (figure 6). DNA is extracted from each cap with the Arcturus PicoPure DNA Extraction Kit (Applied Biosystems) as per protocol F of the PicoPure DNA Extraction Kit user guide.

**Genotyping of DNA extracted from WTS and LCM samples**

A 110 bp region of the human β-globin gene is detected by quantitative real-time PCR to confirm successful DNA extraction, as previously described. DNA from WTS-extracted and LCM-extracted tissue sections are both tested for HPV genotype on the RHA kit HPV SPF10-LiPA25, V.1 (Labo Bio-medical Products BV, Rijswijk, The Netherlands) according to the manufacturer’s instructions. Because of an inability to discriminate between HPV68 and 73 by the SPF10-LiPA25 assay, samples which are positive for HPV68/73 are tested on the Universal Pathogen 200 protocol. The results are reported as positive if there is continuous strong nuclear or nuclear plus cytoplasmic staining of the basal cell layer with extension upwards involving at least one-third of the epithelial thickness. Other patterns are regarded as negative. The second inner section (section 3) is cut to 9 µm thickness and placed onto an Arcturus polyethylene naphthalate (PEN) membrane glass slide (Applied Biosystems, Foster City, California, USA) for LCM and allowed to dry. The next two sections (sections 4 and 5, designated whole tissue sections (WTS)) are cut to 9 µm thickness, placed into sterile 1.5 mL microfuge tubes as scrolls and stored away from light. An additional three sections are cut to 3 µm thickness, mounted and stored for future immunohistochemical staining.

**Digestion and DNA extraction of WTS**

Section 5 is archived at 4°C and section 4 is digested with a mixture of histolene (800 µL) and ethanol (400 µL). The tissue is pelleted at 14 000×g for 10 min, and the resultant pellet washed twice with 500 µL of thickness and automatically stained with CINtec anti-p16 (Roche, Tucson, Arizona, USA) and counterstained with Gills No. 2 haematoxylin (Australian Biostain, Traralgon, Australia) on the Ventana Benchmark Ultra instrument (Ventana Medical Systems, Inc, Oro Valley, Arizona, USA). Sections 1, 2 and 9 are reviewed by a histopathologist experienced in reading biopsy specimens (JMR), to assess p16 positivity and lesion grade. Biopsies are classified as negative for intraepithelial neoplasia, cLSIL, fLSIL, HSIL-AIN2, HSIL-AIN3 or ASCC, and assessment of p16 staining is performed, both in accordance with accepted international criteria. Hence, p16 staining is reported as positive if there is continuous strong nuclear or nuclear plus cytoplasmic staining of the basal cell layer with extension upwards involving at least one-third of the epithelial thickness. Other patterns are regarded as negative. The second inner section (section 3) is cut to 9 µm thickness and placed onto an Arcturus pol...
Conﬁrmed lesions that are positive for β-globin, but not detectable on the SPF10-LiPA25, are tested for HPV DNA on the DNA ELISA kit HPV SPF10, V1 (Labo Bio-medical Products BV) which is validated for detection of 44 HPV genotypes. Samples which are ELISA-positive are genotyped on the RHA Kit HPV SPF+ (Labo Bio-medical Products BV) which is able to identify HPV26, 30, 55, 61, 62, c64, 67, 69, 71, 71sub, 82, 83, 84, 85, 87, 89, 90 and 91. ELISA-negative samples are tested on a qPCR assay targeting the E6 region of HPV16 and 18 as previously described, in addition to HPV45 (forward primer 5’-CTACAAGACGTATCTATTGCC TGTG-3’; reverse primer 5’-CCAGTGTCTCTCATCATATA CAGAGTTT-3’; probe 5’-Cy5-CCTCTGTGCGTTGAGTTGCTT-BHQ3-3’). Conﬁrmed lesions which are still negative on ELISA and/or all previous HPV tests are denoted ‘unas-sessable’. Samples that are positive for β-globin but negative for all HPV tests are denoted ‘HPV not detected’. Samples that are positive for HPV DNA ELISA but not
able to be genotyped are denoted ‘HPV X’. Samples that are negative for mucosal HPV types (negative on SPF10 LiPA, SPF+ LiPA and E6 qPCR) and positive for β-papillomavirus DNA gel electrophoresis assay but not able to be genotyped are denoted ‘HPV β’. Additionally, all WTS samples that are negative for HPV16 or 18 will be tested on a type-specific qPCR assay for these genotypes.

Viral genome methylation sequence analysis
DNA from HPV16-positive HSIL is subjected to methylation sequencing of viral genomic regions associated with regulation of oncogene expression using modifications of published protocols on anal swab, WTS and LCM-extracted tissue. DNA from the second WTS sample (section 5) is extracted using the QIAmp DNA FFPE Tissue Kit (Qiagen GmbH, Hilden, Germany). The same DNA sample used for HPV genotyping of swab and LCM samples is also used for methylation analysis. Known controls for low (SiHa cell line, ATCC HTB35; American Type Culture Collection (ATCC), Manassas, Virginia, USA) and high (CaSki cell line, ATCC CRL1550; ATCC) methylation at the target sites are included in each experiment as well as no-template controls. For each HPV16-positive HSIL, 24 µL of extracted DNA is subjected to bisulfite conversion and DNA cleanup using the Methylamp DNA Modification Kit (Epigentek, Farmingdale, New York, USA), according to the manufacturer’s instructions. This procedure converts unmethylated cytosine (C) residues into thymine (T) residues. Following conversion, two targeted sections of the HPV16 genome are amplified by nested PCR; samples are amplified in a first round of PCR using the primary pair of sequences, and a 1 µL aliquot is then amplified in a second round of PCR using the secondary (nested) PCR primers, as listed in Table 1. The amplified sections are the E2BS1 region from nucleotide position 7348 to 7489 (final amplicon size 142 bp), and E2BS2-4 region from nucleotide position 7839 to 117 (final amplicon size 184 bp). Primary PCR amplifications are performed in reaction mixes with a final volume of 25 µL (12.5 µL of Hot Start Taq Master mix (Qiagen); 0.25 µmol of each primer; 5 µL of bisulfite-modified template DNA; nuclease-free water to final volume). Thermocycler parameters are as follows: 15 min at 95°C, followed by 50 cycles of (40 s at 95°C, 30 s at 50°C and 40 s at 72°C), followed by 6 min at 72°C. Amplicons are visualised on a 2% agarose gel. The secondary (nested) PCR amplification for the E2BS1 region is performed in reaction mixes with a final volume of 25 µL (12.5 µL of Hot Start Taq Master mix (Qiagen); 0.25 µmol of each nested primer; 1 µL of amplicon from the primary PCR; nuclease-free water to final volume). Thermocycler parameters are as follows: 15 min at 95°C, followed by 50 cycles of (40 s at 95°C, 30 s at 56°C and 40 s at 72°C), followed by 6 min at 72°C. The secondary PCR amplification for the E2BS2-4
region is performed in reaction mixes with a final volume of 25 μL (12.5 μL of Hot Start Taq Master mix (Qiagen); 0.25 μmol of each nested primer; 1 μL of amplicon from the primary PCR; nuclease-free water to final volume). Thermocycler parameters are as follows: 15 min at 95°C, followed by 50 cycles of (40 s at 95°C, 30 s at 50°C and 40 s at 72°C), followed by 6 min at 72°C. CpG sequence analysis and quantitation of methylation at nucleotide positions 7426, 7432, 7453, 7459, 7466, 7468, 7486, 7489, 7522 and 7534 on the HPV16 viral genome52 is performed using pyrosequencing technology on the PyroMark Q24 system (Qiagen) at the Australian Genome Research Facility (AGRF). Assay setup, pyrosequence run and analysis are performed by PyroMark Q24 Software. Sequencing primers for E2BS1, E2BS2 and E2BS3,4 regions are listed in Table 1.

Other biomarkers
Additional biomarker analyses may include host genome marker methylation sequencing, viral genome integration frequency and relative transcription levels of HPV mRNA species, contingent on sample availability and assay performance.

Statistical considerations
In total, about 600 participants (30–40% of whom will be HIV-positive) will be recruited by June 2015. As of the end of 2014, 480 participants have been recruited with 125 (26%, 95% CI 22% to 30%) men diagnosed with biopsy-confirmed LSIL and 143 (30%, 95% CI 26% to 34%) men with biopsy-confirmed anal HSIL. Therefore, we anticipate between 130 and 180 of the total 600 participants will be diagnosed with LSIL, and between 150 and 200 will be diagnosed with biopsy-confirmed HSIL at baseline.

HPV prevalence in LCM-isolated lesions, including low-risk (including β-papillomavirus or HPV X), high-risk and individual genotypes will be calculated overall and stratified by HIV status as the proportion of lesions that tested positive for a specific HPV type or risk group. The exact binomial method will be used to calculate the corresponding 95% CIs for prevalence values. Biomarker positivity associated with lesion type will be calculated using the same method. Univariate and multivariate logistic regression models will be carried out to measure the strength of association between HPV status and each biomarker in lesions at baseline.

To determine the performance of biomarkers in predicting disease persistence at 12 months, univariate and multivariate Cox regressions will be performed to assess the strength of the association between biomarker positivity and anal HSIL persistence at 12 months in those who are diagnosed with anal HSIL at baseline. The sensitivity and specificity of each biomarker detected on anal swab to predict persistent HSIL at 12 months will be calculated alone and in combination with HPV genotyping results. In addition, univariate logistic regression models will be carried out to measure the strength of each test in relation to a diagnosis of HSIL. A multivariate logistic regression model will be constructed to

### Table 1

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2BS1 amplicon</td>
<td>7348 5’-Biotin-ATTGTGGTTGTTGTTATTATTGTA-3’</td>
<td>175</td>
</tr>
<tr>
<td>E2BS1 F2 bio</td>
<td>7371</td>
<td></td>
</tr>
<tr>
<td>E2BS1 reverse primer Ampy</td>
<td>7522 5’-AACCATAATTACTAACATAAAACT-3’</td>
<td>7499</td>
</tr>
<tr>
<td>E2BS1 R3</td>
<td>7371</td>
<td></td>
</tr>
<tr>
<td>Nested amplification forward primer Ampy</td>
<td>7348 5’-Biotin-ATTGTGGTTGTTGTTATTATTGTA-3’</td>
<td>142</td>
</tr>
<tr>
<td>Nested amplification reverse primer Ampy</td>
<td>7489 5’-AACACATTATTACAAAAAC-3’</td>
<td>7468</td>
</tr>
<tr>
<td>Sequenting primer</td>
<td>7486 5’-ACATTATTACAAAAACAC-3’</td>
<td>7466</td>
</tr>
<tr>
<td>Sempy E2BS1 R2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>E2BS2-4 amplicon</td>
<td>7830 5’-TTGTAANATGTTATGTTGTTGTA-3’</td>
<td>193</td>
</tr>
<tr>
<td>E2BS2,3,4 F</td>
<td>117 5’-Biotin-AAATCTAAAAACATTACATTCTC-3’</td>
<td>94</td>
</tr>
<tr>
<td>E2BS2,3,4 R bio</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Nested amplification forward primer Ampy E2BS2 F</td>
<td>7839 5’-TGATATATGGGTGTGTTAAT-3’</td>
<td>7859</td>
</tr>
<tr>
<td>Nested amplification reverse primer Ampy E2BS2 R bio</td>
<td>117 5’-Biotin-AAATCTAAAAACATTACATTCTC-3’</td>
<td>94</td>
</tr>
<tr>
<td>Sequencing primer</td>
<td>7839 5’-TGATATATGGGTGTGTTAAT-3’</td>
<td>7859</td>
</tr>
<tr>
<td>Sempy E2BS2 F</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Sequencing primer</td>
<td>9 5’-AATTATGGTTATAAAATTAGGGG-3’</td>
<td>30</td>
</tr>
</tbody>
</table>

determine a combination of these biomarkers that best predict persistent HSIL and to assess the association with HIV status and other covariates such as age.

Percentages (%) agreement and $\kappa$ statistic will be used to determine HPV concordance in anal swab, WTS and LCM sample (poor if $\kappa \leq 0.20$; fair if $0.21 \leq \kappa \leq 0.40$; moderate if $0.41 \leq \kappa \leq 0.60$; substantial if $0.61 \leq \kappa \leq 0.80$; good if $\kappa > 0.80$). Differences in HPV genotypes prevalence between samples will be assessed by the McNemar test.

Currently available algorithms to determine the most likely lesion-associated genotype based on swab samples will be compared with actual lesion-associated genotypes determined using LCM. Percentage detection and 95% CIs will be used to calculate the performance of hierarchical and proportional attribution methods against LCM-determined lesion-associated genotypes.

Ethics and dissemination

Written informed consent is obtained from all individuals before any study-specific procedures are performed. As clinical guidelines do not indicate treatment and there is currently no scientific evidence to support routine treatment of anal HSIL, participants diagnosed with HSIL are not routinely treated, except where deemed clinically necessary; this may include severely abnormal blood vessels, persistent large lesions or other risk factors for ASCC, such as prolonged immunosuppression, low CD4 nadir and history of smoking.

Findings from this study will be disseminated to participants and the community through study newsletters, and through peer-reviewed publications and international conferences.

DISCUSSION

Importance of this study

With continuing increases in ASCC incidence, future prevention of this disease will rely on effective vaccination and screening measures and these in turn will depend on accurate data pertaining to the association of individual HPV genotypes with ASCC and with anal HSIL. Persistent anal HSIL is widely accepted to be the precursor to ASCC, and is the anal counterpart to cervical HSIL, the main disease end point in cervical screening programmes. The end point for anal screening will likely differ from cervical screening as rates of progression from anal HPV infection and/or HSIL to anal cancer are very low, and the populations that are likely to be most frequently screened contain a high proportion of individuals with HR-HPV and/or HSIL. Effective anal screening programmes will therefore require very high specificity for an appropriate end point, to limit unnecessary procedures on a large number of low-risk individuals, as well as high sensitivity to identify the small population of high-risk individuals. This research will contribute to the determination of appropriate end points for anal screening.

Approximately 70–90% of cases of ASCC are attributed to infection with HPV16. However, the exact contribution of HPV16, and perhaps more importantly other HPV genotypes, to the load of ASCC cases is not known, as up to a quarter of anal cancer specimens test positive for multiple HPV genotypes. Where HPV16 is present, it is usually assumed to be the causative genotype; given that HPV16 is a commonly found genotype in any population of sexually active adults, it is highly likely to be incidental to the cancer in a number of cases and therefore its contribution to cancer development overstated. Conversely, an increasing quantity of evidence suggests that so-called LR-HPV genotypes—particularly HPV6 and 11—contribute to a larger proportion than expected based on cervical cancer data sets and present data collection methods probably lead to underestimation of their contribution to anal cancer rates.

In addition to HPV genotype detection, methylation patterns also have potential utility as biomarkers in cancer screening, particularly with rapid advances in sequencing technology and the increasing recognition in other medical fields of the diagnostic potential of epigenetic markers. Currently available data on methylation of viral DNA targets are insufficient to inform development of methylation as a screening marker for HPV-related cancers, and are derived predominantly from cervical tissue and cervical cell lines with only a single study presenting any methylation sequence data for the HPV genome in anal lesions.

Because it is the persistence of HR-HPV and HSIL that lead to increased risk of anal cancer development, longitudinal studies provide more useful information than cross-sectional data, particularly for rare events such as anal cancer incidence. Longitudinal or natural history studies provide data to allow more accurate risk categories to be assigned to single screening results, so that individual patients can be triaged appropriately. This prospective, longitudinal study is designed to provide data on risk factors associated with persistent HSIL—and by extension increased risk of anal cancer development—to inform future decisions on anal cancer prevention in at-risk populations.

Strengths of this study

This participant cohort is a unique resource, as one of only a very small number of cohorts with prospectively collected longitudinal cytology and histology samples from HIV-positive and HIV-negative participants diagnosed with the full spectrum of HPV-associated anal lesions. A team of intensively trained anoscopists have their diagnostic skills monitored regularly with extensive quality control metrics, ensuring high-quality clinical data. Participants in this cohort are not routinely treated when HSIL is detected unless it is determined to be clinically necessary, for example, severely abnormal blood vessels, persistent large lesions or other risk factors for ASCC, such as prolonged immunosuppression, low CD4
nadir and history of smoking. This approach allows for longitudinal follow-up of lesions, and therefore allows assessment of when treatment is likely to be most beneficial.

A major strength of the present study is the use of LCM to distinguish lesions from adjacent normal tissue, which allows for precise genotype and biomarker testing from which to infer the natural history of HPV-associated anal lesions. Methylation patterns and the behaviour of other biomarkers are also often tissue-specific or lesion-specific, and can be heterogeneous within even a small tissue sample. The use of LCM generates extremely low-concentration DNA samples, and the assays being used in this study are very sensitive for low copies of DNA targets, and have a small amplicon size for the detection of fragmented DNA, which is partially degraded in the fixation process. These methods have been proven sensitive enough for HPV genotyping and methylation sequence analysis in published studies and in our own laboratory.

In routine diagnostic work, p16 immunostaining is utilised, as per LAST criteria, for confirming a diagnosis of HSIL-AIN2 and for distinguishing true HSIL-AIN3 from atypical immature metaplasia. This latter differential diagnosis is more challenging in the anal canal that in the cervix (T Darragh, personal communication, 2013). Routine p16 staining is not indicated for diagnostic samples due to the uncertain significance of p16-positive LSIL, but it is a useful research tool for further exploring the natural history of p16-positive LSIL in conjunction with HPV genotype and/or biomarker analysis and potentially clarifying the significance of these lesions.

Anticipated challenges
The major challenges of this study relate to diagnostic uncertainty and establishing whether or not a lesion is persistent from one visit to the next. It is inevitable that some lesions will be missed during HRA, and therefore multiple sources of data will be used to define the persistence of individual lesions, as indicated in figure 2. The study uses both cytology and histology (based on HRA-guided biopsy collection) to establish disease grade in participants at each visit. It is known that anal cytology and HRA/histology each have their limitations, and therefore both methods will be used in a complementary fashion to derive a more accurate diagnosis. Longitudinal sampling and identification of the same individual lesion has not been attempted before in a systematic manner. This adds an extra dimension of complexity to the project. Defining persistent lesions will not be straightforward due to the difficulties inherent in sampling the anal canal, the possibility of missed lesions, the probability of multiple lesions and the uncertainties in locating and identifying the same lesion at consecutive visits. The presence of multiple distinct lesions caused by the same genotype cannot be ruled out. All anoscopists are aware of each participant’s previous clinic results and have access to HRA photographs; this improves the likelihood of sampling the same lesion during subsequent HRA visits. It is known that the anal canal may rotate up to 90° between clinic visits (personal communication R J Hillman, 2012), making precise location of anatomical features difficult; this has been acknowledged and follow-up biopsies will be considered if they are taken up to one octant either side of the baseline biopsy to reduce the chance of missing a persistent lesion. The use of features in HRA photographs will also aid in determining lesion location, and in conjunction with anoscopists’ notes will be used to indicate the extent of disease (ie, across multiple octants). Each original biopsy diagnosis is made by one of a team of three experienced anogenital pathologists, and this is the basis on which tissues are selected for inclusion. The review diagnoses, however, are all made by one of these three pathologists (JMR) based on the new H&E and p16 slides as described in the Methods section. The main issue resulting from this will be confirming persistence of HSIL when the review diagnosis at follow-up has been downgraded to LSIL or negative for SIL. In these cases, the study investigators will make a decision based on the original diagnosis, the HPV genotype detected by LCM, and the histological diagnosis at the next clinic visit.

This study will use precise molecular biology techniques and international expertise to characterise the natural history and biological markers of individual anal lesions in a population at high risk of anal cancer. This is an ambitious and unique study which will enhance our understanding of the clinical and biological behaviour of HPV-related anal lesions and the potential prognostic implications of future HPV genotype or biomarker screening tests.

Author affiliations
1Regional HPV Labnet Reference Laboratory, Department of Microbiology and Infectious Diseases, The Royal Women’s Hospital, Parkville, Victoria, Australia
2Murdoch Children’s Research Institute, Parkville, Victoria, Australia
3Douglas Hanly Moir Pathology, Sydney, New South Wales, Australia
4Western Sydney Sexual Health Centre, University of Sydney, Sydney, New South Wales, Australia
5The Kirby Institute, UNSW Australia, Kensington, New South Wales, Australia
6RPA Sexual Health, Sydney Local Health District, Camperdown, New South Wales, Australia
7Department of Obstetrics and Gynaecology, University of Melbourne

Collaborators The SPANC study team includes: AEG, IMP, FJ, Brian Acraman, Garrett Prestage, Leonie Crampton, Patrick McGrath, Kathy Petoumenos, Matthew O’Dwyer, Robert Mellor, Piero Pezzopane, Matthew Law (Kirby Institute, University of New South Wales); Carmella Law, Winnie Tong, Daniel Seeds, Eddie Fraissard and Andrew Carr (St Vincent’s Hospital, Sydney); DJT and Rick Varma (RPA Sexual Health, Sydney); Kit Fairley (Melbourne Sexual Health Centre); RJH, Kirsten McCaffrey and Kirsten Howard (University of Sydney); Annabelle Farnsworth, JMR, Adele Richards and Julia Thorloe (Douglas Hanly Moir Pathology, Sydney); SMG, SNT, AMP and SP (The Royal Women’s Hospital, Melbourne); and Lance Feeney and Russ Glynn (community representatives).

Contributors All authors contributed to study design. AMC drafted the manuscript. JMR provided histopathological expertise. SP and MM provided molecular biology expertise. DMM provided statistical expertise. RJH provided clinical expertise. AEG, FJ, IMP, DJT, SNT collectively conceived of
the study and provided epidemiological and molecular biology expertise. All authors read and approved of the final manuscript.

Funding The SPANC study is funded by Australian Commonwealth Government National Health and Medical Research Grant programs (#686971 and #1071289) and a Cancer Council New South Wales Strategic Research Partnership Program grant (#13-11).

Competing interests None declared.

Ethics approval The St Vincent’s Hospital Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES


