AIDS, Publish Ahead of Print

DOI: 10.1097/QAD.0000000000002381

Human papillomavirus infection and cervical dysplasia in HIV-positive women: potential role of the vaginal microbiota

Janneke H.H.M. VAN DE WIJGERT[1,2*], A. Christina GILL[1*†], Admire CHIKANDIWA[3], Marijn C. VERWIJS[1], Helen A. KELLY[4], Tanvier OMAR[5], Sinead DELANY-MORETLWE[3], Michel SEGONDY[6], Suzanna FRANCIS[4], Alistair C. DARBY[7], Philippe MAYAUD[4], for the HARP Study Group‡

Affiliations:
1. Institute of Infection and Global Health, University of Liverpool, Ronald Ross building, 8 West Derby Street, Liverpool L69 7BE, United Kingdom
2. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Stratenum building, 3508 GA Utrecht, The Netherlands
3. Wits Reproductive Health & HIV Institute, Faculty of Health Sciences, University of the Witwatersrand, Hillbrow Health Precinct, 22 Esselen Street, Hillbrow, 2001 Johannesburg, South Africa
4. London School of Hygiene and Tropical Medicine, Keppel Street, Bloomsbury, London WC1E 7HT, United Kingdom
5. National Health Laboratory Services, 1 Modderfontein Road, 2131 Sandringham, Johannesburg, South Africa
6. Pathogenesis and Control of Chronic Infections, INSERM, University of Montpellier, 60 Rue de Navacelles, 34394 Montpellier, France
7. Centre for Genomic Research, University of Liverpool, Biosciences building, Crown Street, Liverpool L69 7ZB, United Kingdom

* Shared first authors
† Current affiliation: Biomedical Services, University of Oxford, Oxford, United Kingdom
‡ HARP Study Group composition in the acknowledgements

Correspondence:
Professor Janneke van de Wijgert
Institute of Infection and Global Health, University of Liverpool
Ronald Ross building, 8 West Derby Street, Liverpool L69 7BE, United Kingdom
ORCID: 0000-0003-2728-4560
E-mail: j.vandewijgert@liverpool.ac.uk
Mobile phone: +44-7557195108

Running head: Vaginal microbiota and HPV dysplasia.

Funding: European Commission 7th Framework Programme [original HARP study; grant agreement number HEALTH-2010-F2-265396] and the University of Liverpool [16S rRNA sequencing].

ABSTRACT

Objectives: To assess the associations between microbiological markers of vaginal dysbiosis and 1) incident/cleared/type-swap/persistent high-risk human papillomavirus (hrHPV) infection; and 2) incident/cured/cleared/persistent high-grade cervical intraepithelial neoplasia (CIN2+) while controlling for persistent hrHPV infection.

Design: Two nested case-control studies (N=304 and 236) within a prospective cohort of HIV-positive women in Johannesburg, South Africa.

Methods: Participants were examined for hrHPV type (InnoLipA), cervical dysplasia (histology), and vaginal microbiota (VMB) composition (V3-V4 Illumina HiSeq 2x300bp) at baseline and endline, a median of 16 months later.

Results: Women with incident hrHPV compared to those who remained hrHPV-negative were less likely to have an optimal *Lactobacillus crispatus*/*jensenii*-dominated VMB type at endline (relative risk ratio (RRR)=0.125, p=0.019) but not at baseline. Having different hrHPV types at both visits was associated with multiple anaerobic dysbiosis markers at baseline (e.g. increased BV-anaerobes relative abundance: RRR=3.246, p=0.026). Compared to women without CIN2+ but with hrHPV at both visits, women with incident CIN2+ had increased Simpson diversity (RRR=7.352, p=0.028) and non-significant trends in other anaerobic dysbiosis markers at endline but not baseline. These associations persisted after controlling for age, hormonal contraception, and CD4+ count. Current hormonal contraceptive use (predominantly progestin-only injectables) was associated with increased CIN2+ risk over-and-above persistent hrHPV infection and independent of VMB composition.

Conclusions: hrHPV infection (and/or increased sexual risk-taking) may cause anaerobic vaginal dysbiosis, but a bidirectional relationship is also possible. In this population, dysbiosis did not increase CIN2+ risk, but CIN2+ increased dysbiosis risk. The CIN2+ risk associated with progestin-only injectable use requires further evaluation.

Keywords: HPV, cervical intraepithelial neoplasia (CIN), cervical cancer, vaginal microbiota, lactobacilli, vaginal dysbiosis, 16S rRNA gene sequencing, women, HIV, South Africa.

INTRODUCTION

Cervical cancer is the fourth most common cancer affecting women worldwide [1]. Persistent genital infection with high-risk human papillomavirus (hrHPV) types is a necessary trigger, but hrHPV infection and high-grade cervical intraepithelial neoplasia (CIN) lesions may regress without treatment [2]. Cervical cancer and hrHPV prevalences are particularly high in sub-Saharan Africa, including South Africa, which currently also has the highest HIV prevalence worldwide [3]. Women living with HIV have a higher prevalence of genital hrHPV infection than the general population [4], and a higher risk of progression to CIN or cervical cancer, likely due to HIV-induced immunosuppression [5].

Persistent hrHPV infection does not always result in CIN/cancer, and other exposures are thought to play important roles. One of these is vaginal microbiota (VMB) dysbiosis. An optimal VMB is dominated by lactobacilli. The most common type of vaginal dysbiosis is bacterial vaginosis (BV), which is characterised by a persistent decrease in lactobacilli and increase in fastidious anaerobes (referred to as BV-anaerobes). BV affects 30-40% of women worldwide at any given time [6]. A recent systematic review and meta-analysis of 14 longitudinal studies published between 2003 and 2017 showed that BV is associated with increased risks of incident hrHPV (relative risk 1.33), hrHPV persistence (1.18), and CIN/cancer (2.01) [7]. However, most of these studies (11/14) used crude VMB assessments by microscopic methods only. Furthermore, in all studies, women with high-grade CIN/cancer were compared with those without, regardless of their hrHPV status. The review could therefore not disentangle the impact of VMB characteristics on hrHPV infection from progression to CIN/cancer during or after persistent hrHPV infection.

We addressed the shortcomings of these earlier studies by conducting two nested case-control studies within the 'HPV in Africa Research Partnership' (HARP) study in South African women living with HIV [8,9]: the hrHPV sub-study and the CIN sub-study. We used appropriate control groups in each sub-study and incorporated molecular VMB assessments.

METHODS

The HARP study aim was to improve cervical cancer prevention programs for HIV–infected African women and procedures have been described elsewhere [8,9]. HIV-infected women (N=623), aged 25-50, were recruited from HIV treatment centres and surrounding communities in Johannesburg in 2011 and 2012. Exclusion criteria were previous treatment for cervical cancer, previous hysterectomy, and being pregnant or less than eight weeks postpartum. Enrolment was stratified by antiretroviral therapy status (on therapy versus therapy-naïve, following the 2010 World Health Organisation guidelines that used a CD4+ 350 cells/μl cut-off [10]) in a 2:1 ratio. Women were followed up every six months for a median of 16 months (range 11-22 months), but only baseline and endline data were used for the VMB sub-studies. The hrHPV sub-study (N=304) included women with CIN1 or lower (≤CIN1) at both visits who had never received cervical treatment, and categorised them into mutually exclusive categories as follows (Figure 1): twice hrHPV-negative (n=37; referred to as persistent hrHPV-negative controls), incident hrHPV (n=43; no hrHPV types at baseline but at least one at endline), cleared hrHPV (n=65; at least one hrHPV type at baseline and none at endline), hrHPV type-swap (n=67; different hrHPV types at the two visits), and persistent type-specific hrHPV (n=92). The CIN sub-study (N=236) included women with CIN2 or higher (CIN2+) on at least one visit, who did or did not receive cervical treatment in between the two visits, and categorised them into mutually exclusive categories as follows: incident CIN2+ (n=22), cured CIN2+ after treatment (n=50), spontaneous CIN2+ clearance

in the absence of treatment (n=14), persistent or recurrent CIN2+ (n=25; five women were treated but had CIN2+ recurrence and 20 women were not treated and had CIN2+ at both visits), and prevalent CIN2+ (n=33; data available for one visit only). In this sub-study, women with ≤CIN1 but persistent type-specific hrHPV at both visits were used as the control group (n=92; referred to as persistent hrHPV-positive controls). This group was included in both sub-studies, as a comparison group in the hrHPV sub-study and as a control group in the CIN sub-study, making the total sample size of women enrolled in the two sub-studies 448.

Diagnostic laboratory assessments

HIV-1 serostatus was diagnosed according to national guidelines [11]. CD4+ T-lymphocyte counts were determined by FACScount (Becton-Dickinson, Franklin Lakes, New Jersey, USA) and plasma HIV viral load by COBAS Taqman (Roche Diagnostics, Johannesburg, South Africa). All samples were assessed for hrHPV qualitatively by Digene HC-II or CareHPV test (both Qiagen, Gaithersburg, Maryland, USA) and by INNO-LiPA HPV Genotyping Extra (Fujirebio, Courtaboeuf, France), classifying 13 types as high-risk (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68; Supplement 1: Methods, http://links.lww.com/QAD/B541) [12]. Cervical dysplasia status was assessed by Papanicolaou smear cytology, visual inspection with acetic acid or Lugol's iodine, and colposcopy. If any of these or qualitative hrHPV testing were positive, histology of four-quadrant cervical biopsies was done using the three-tier CIN classification system [13], and binarised as ≤CIN1 or CIN2+ based on the highest reading. All histological slides from women with CIN2+ and 5-10% of slides from women with ≤CIN1 were reviewed by the HARP Endpoint Committee of five pathologists for consensus classification [14]. BV was diagnosed by Gram stain Nugent scoring [15], and vaginal yeast infection by the presence of yeasts on Gram stain. Women were also screened for other sexually transmitted infections

(STIs; Supplement 1: Methods, http://links.lww.com/QAD/B541). hrHPV, cervical dysplasia, and CD4+ count assessments were done at both visits, and the other procedures at baseline only.

Molecular VMB testing

Vaginal swabs were stored at room temperature in Boonfix medium (a patented fixative containing ethanol, low molecular weight polyethylene glycol, and acetic acid). We first determined that this fixative was suitable for sequencing (Supplement 1: Methods, http://links.lww.com/QAD/B541). DNA was extracted from one swab per woman per visit (N=873), using lysozyme lysis combined with the Qiagen DNeasy Blood and Tissue kit (Qiagen, Manchester, UK) with inclusion of the swab head up to and including the proteinase K step. The 16S rRNA gene V3-V4 region was amplified and sequenced on an Illumina HiSeq 2500 instrument (Illumina, San Diego, CA, USA).

Molecular data processing

Molecular data processing steps are described in Supplement 1 (Methods, http://links.lww.com/QAD/B541). Briefly, Swarm v2.1.13 was used to assign reads to operational taxonomic units (OTUs) and taxonomic assignments were made using RDP classifier against the Silva v128 database (Supplement 1: Methods, http://links.lww.com/QAD/B541) [16]. Low read count (<100 reads), non-bacterial, and contaminant OTUs were removed. Relative abundances were rarefied at 1,039 reads using the GUniFrac v1.0 package in R v3.4.1 (R Foundation for Statistical Computing 2015, Vienna, Austria). The rarefied OTU relative abundance table consisted of 1981 OTUs in 871 samples, of which 246 were non-minority OTUs (relative abundance of at least 1% in at least one sample).

Data reduction for biostatistical modelling was done in three different ways. First, the Simpson diversity index (1-D) was calculated for each sample, ranging from zero (no diversity) to one (infinite diversity). Second, each OTU was assigned to one of four 'bacterial groups' based on the published literature (Supplement 2, http://links.lww.com/QAD/B541) as follows: 1) lactobacilli; 2) BV-anaerobes; 3) pathobionts (bacteria with higher intrinsic pathogenicity than BV-anaerobes that are not typically associated with BV; also includes STI pathogens); and 4) 'other bacteria' (skin and Bifidobacteria). The proportion of sequencing reads assigned to a specific OTU in a sample is called the relative abundance of that OTU. For each sample, relative abundances of OTUs belonging to the same bacterial group were summed (Supplement 1: Methods, http://links.lww.com/QAD/B541). This resulted in four relative abundances (one for each bacterial group) per sample, which sum to one. Third, we used hierarchical clustering based on Euclidean distance to pool samples into seven VMB types (Supplement 1: Figure S1, http://links.lww.com/QAD/B541). 1) *Lactobacillus iners*-dominated (Li; >70% lactobacilli of which *L. iners* was the most common); 2) *L. crispatus* or *L. jensenii*-dominated (Lcj; also >70% lactobacilli of which *L. crispatus* or *L. jensenii* were the most common); 3) lactobacilli and anaerobes (LA; >10% lactobacilli with the remainder BV-anaerobes); 4) high diversity BV-anaerobes (BV), 5) BV-anaerobe-dominated (AD; >50% *Gardnerella vaginalis* or *Atopobium vaginae*); 6) pathobionts-characterized (PB; >25% pathobionts) and 7) *Bifidobacterium*-dominated (BD). An increased Nugent score, Simpson index, or BV-anaerobes relative abundance, a reduced lactobacilli relative abundance, and VMB types LA, BV, and AD were considered markers of (transition to/from) anaerobic vaginal dysbiosis. The pathobionts and 'other bacteria' bacterial groups, and PB and BD VMB types, represent poorly characterised VMB states and were uncommon.

Statistical analyses

Statistical analyses were performed in R v3.4.1 and Stata v13 (StataCorp, College Station, TX, USA). Baseline characteristics between the parent study and sub-studies were compared by Kruskal Wallis test for continuous and Chi-squared or Fisher's exact test for categorical variables. Unadjusted multinomial logistic regression models were carried out for three multi-category outcome variables: the hrHPV and CIN outcomes as described above and in Figure 1, and a combined hrHPV/CIN outcome (4 mutually exclusive groups; N=448) to optimise statistical power: ≤CIN1 and no hrHPV at both visits (n=37), ≤CIN1 at both visits and hrHPV at one visit (n=108), ≤CIN1 and hrHPV at both visits (n=159), and CIN2+ regardless of hrHPV status at one or both visits (n=144). Each unadjusted model included one outcome and one baseline or endline VMB variable but yielded multiple relative risk ratios (RRRs) due to the multi-category nature of the outcomes. For example, models with the combined outcome yielded three RRRs: case group 1 versus controls, case group 2 versus controls, and case group 3 versus controls. Multivariable models were carried out in the same manner but for the combined outcome only, and were adjusted for age, CD4+ count at baseline or endline, and hormonal contraceptive use at baseline or during the study. Potential confounders were selected a priori based on the published literature and the above-mentioned confounders were consistently associated with hrHPV/CIN outcomes and VMB variables in our data.

Ethical statement

All participants provided written informed consent. The study was conducted in accordance with the Helsinki Declaration, and approved by the research ethics committees of Witwatersrand University, London School for Hygiene and Tropical Medicine, and University of Liverpool.

RESULTS

The 448 women selected for the sub-studies did not differ from those not selected (N=175), except that they were more likely to be on antiretroviral therapy (70.1% versus 52.6%) (Supplement 1: Table S1, http://links.lww.com/QAD/B541). The median age was 34 (interquartile range 30-39) (Table 1). At baseline, the majority of women reported to have a regular male sex partner and one sex partner in the last three months. Vaginal cleansing was practiced at least weekly by 41.5% of women. A quarter (25.2%) reported current hormonal contraceptive use, which consisted predominantly of intramuscular injections of medroxyprogesterone acetate or norethindrone enanthate. The median CD4+ count was 423 (interquartile range 317-566). STIs other than HIV and HPV were common, especially herpes simplex virus type 2 (95.3%) and *Trichomonas vaginalis* (15.6%), but vaginal yeast infections were less common (7.5%). None of these characteristics differed between the hrHPV and CIN comparison groups, except for injectable contraceptive use, which was significantly more common in the cured, cleared, and persistent CIN2+ groups (34.0, 57.1, and 36.0%) compared to the other groups (4.6-24.3%) (Supplement 1: Table S2, http://links.lww.com/QAD/B541).

The majority of women had at least one hrHPV type at baseline (79.7%) and endline (70.8%), and 42.9% had BV by Nugent score at baseline. The mean Simpson index was 0.54, and the mean bacterial group relative abundances 0.48 for lactobacilli, 0.49 for BV-anaerobes, 0.03 for pathobionts, and 0.01 for 'other bacteria', for all women combined at baseline (Table 1). The most common VMB types at baseline were Li (26.3% of women), LA (25.3%), and BV (33.8%) (Table 1). The Lcj (6.2%) and AD (6.6%) types were less common, and the PB and BD types were rare. As expected, the median Simpson index and mean bacterial group relative abundances differed per VMB type (Supplement 1: Figure S1, http://links.lww.com/QAD/B541). The overall proportions of women with each VMB type

were similar at endline, but only 143/414 (35%) of individual women with two VMB assessments had the same VMB type at both visits. Women with the Lcj or BV type at baseline were most likely to have the same VMB type at endline (50.7% and 44.4%, respectively) while those with Li (25.5%), LA (26.2%), or AD types (14.8%) were less stable. The median Bray-Curtis similarities between baseline and endline samples were: Lcj 33%, Li 44%, LA 40%, BV 22%, BD 19%, and PB 6.7%.

Stacked bar graphs of the proportions of women with each VMB type at baseline and endline by outcome group are shown in Figure 2, and data for all VMB characteristics by outcome group in Supplement 1 (Table S3, http://links.lww.com/QAD/B541). In the hrHPV sub-study, unadjusted multinomial logistic regression models showed that baseline VMB compositions of women who acquired or cleared hrHPV during the study did not significantly differ from the baseline VMB compositions of the persistent hrHPV-negative controls (Table 2). However, women who acquired hrHPV were much less likely to have an optimal Lcj VMB type at endline (RRR=0.125, p=0.019), and women who cleared hrHPV were more likely to have anaerobic dysbiosis markers at endline (reaching significance for Simpson diversity: RRR=3.856, p=0.034). Women in the type-swap group were more likely to have anaerobic dysbiosis markers at baseline (significantly increased Nugent score, Simpson diversity, and BV-anaerobes relative abundance, and significantly decreased Lactobacilli relative abundance), and women in the type-swap and type-specific persistence groups had non-significant trends towards a lower likelihood of having an optimal Lcj VMB type at endline (Table 2).

In the CIN sub-study, unadjusted multinomial logistic regression models showed that women who had CIN2+ at least once during the study were more likely to have anaerobic dysbiosis markers at baseline and endline than persistently hrHPV-negative controls (Table 2).

However, when compared to persistent hrHPV-positive controls, the endline – but not the baseline – VMB compositions of women with incident or persistent/recurrent CIN2+ were more dysbiotic.

The impact of potential confounders on these associations was assessed in multinomial logistic regression models with the combined hrHPV/CIN outcome (Table 3). Women who had CIN2+ at least once during the study, compared to persistently hrHPV-positive controls, were younger (RRR=0.954; p=0.021), more likely to use hormonal contraception at baseline (RRR 2.358; p=0.001) and during the study (RRR=1.317; p=0.008), and had a higher $\log_{10}$ HIV-1 plasma viral load (RRR=1.200, p=0.050). These same women compared to persistently hrHPV-negative controls were younger (RRR=0.881; p<0.001) and had a trend towards lower CD4+ count (RRR=0.998, p=0.053). Smoking, sexual behaviour, vaginal cleansing, bacterial/protozoal STIs, and yeasts on vaginal Gram stain were not significantly associated with any of the outcomes. The associations between the combined hrHPV/CIN outcome and individual VMB composition variables one at a time were similar to those in Table 2 for the more detailed separate hrHPV and CIN outcomes: hrHPV at one visit compared to persistent hrHPV-negative controls was associated with none of the VMB variables, hrHPV at both visits compared to persistent hrHPV-negative controls was associated with baseline anaerobic dysbiosis, and CIN2+ at least once compared to persistent hrHPV-positive controls was associated with endline anaerobic dysbiosis (Table 3). In addition, CIN2+ at least once compared to persistent hrHPV-negative controls was associated with anaerobic dysbiosis at baseline and endline. All multivariable models were controlled for age, hormonal contraceptive use, and CD4+ count at baseline or endline, but this did not change the results (Table 3).

DISCUSSION

Our results support the associations between vaginal dysbiosis, hrHPV, and high-grade CIN that have been reported previously [7] and are most consistent with changes in hrHPV infection and CIN status preceding changes in VMB composition. However, we compared data from two visits only, about 16 months apart, and bidirectional relationships can therefore not be ruled out. Women who did not have any hrHPV at baseline and acquired hrHPV during follow-up were similar to persistent hrHPV-negative controls at baseline but not at endline, suggesting that hrHPV acquisition altered the VMB rather than vice-versa. Women with hrHPV type-swap were more dysbiotic at baseline than women with persistent type-specific hrHPV infection, suggesting that frequency of hrHPV acquisition may determine the likelihood of anaerobic dysbiosis to a larger extent than type-specific hrHPV persistence. However, women who have increased exposure to hrHPV types most likely also have exposure to a higher number of male sex partners and/or sex acts, which in themselves are risk factors for vaginal dysbiosis [17]. We assessed multiple sexual behaviour characteristics for potential confounding but residual confounding may have been present.

Women with incident CIN2+ were more likely to have anaerobic dysbiosis at endline, but not baseline, when compared to women with hrHPV infection but no CIN2+ at both visits. Anaerobic dysbiosis risk therefore seems to increase concurrently with CIN2+ development, suggesting that VMB composition does not play a role in CIN2+ development over-and-above the presence of a persistent hrHPV infection. CIN2+ at least once compared to persistent hrHPV-negative controls was associated with anaerobic dysbiosis at baseline and endline, which is similar to findings from previous studies to date that compared women with and without CIN2+ regardless of hrHPV status [7]. hrHPV status should therefore be taken into account when evaluating the VMB-CIN relationship, as we had hypothesised.

An additional important finding was that hormonal contraceptive use (which consisted predominantly of intramuscular injections of medroxyprogesterone acetate or norethindrone enanthate) increased CIN2+ risk over-and-above persistent hrHPV infection. Hormonal contraception is known to decrease anaerobic dysbiosis risk [18], and current use was associated with decreased likelihood of anaerobic dysbiosis in this study (data not shown). The association with CIN2+ risk is therefore unlikely to be mediated by VMB composition. In the overall South African HARP cohort of 623 women, CIN2+ prevalence at baseline was higher among current injectable users compared to never or past users of any hormonal contraception (odds ratio 2.75, adjusted for condom use, antiretroviral therapy status, and CD4+ count) [19]. No association was observed for CIN2+ prevalence and combined oral contraceptive use. The potential roles of different progestin-only injectable contraceptives on CIN2+ development by hrHPV and HIV status should be evaluated urgently.

Only three other studies have assessed associations between molecular VMB assessments and hrHPV outcomes longitudinally, and none for CIN outcomes. Brotman et al reported that lactobacilli-dominance, and particularly *L. gasseri*-dominance, was associated with decreased HPV incidence and increased HPV remission rates in 32 HIV-negative women [20]. Reimers et al (N=68) found that high *L. crispatus* relative abundance reduced HPV detection during follow-up (after controlling for HIV-status), but found no associations with lactobacilli as a group or for other individual *Lactobacillus* species [21]. Di Paola et al compared baseline samples from 27 women who had cleared their HPV infection one year later with 28 women with persistent infection (all HIV-negative), and found a difference in anaerobic dysbiosis prevalence (7.4% and 43.0%, respectively) [22]. None of these studies contradict our findings, but they do suggest that a bidirectional relationship between VMB and HPV status is likely.

HPV causes alterations in cell physiology as well as innate immune response suppression of infected cervicovaginal mucosal cells [23]. Neoplastic cells have a drastically increased glucose demand compared to healthy cells, and ferment glucose into lactate instead of carbon dioxide. These altered mucosal environments likely facilitate BV-anaerobe growth at the expense of *Lactobacillus* growth which causes cervicovaginal dysbiosis [24]. However, cervicovaginal dysbiotic states (which could be caused by multiple factors in addition to HPV infection or neoplastic cells) reduce cervicovaginal barrier function [25] and alter metabolic profiles [24], and these may in turn facilitate HPV acquisition and CIN/cancer development, respectively.

Strengths of our study included a much larger sample size than similar previous studies, high quality longitudinal hrHPV, CIN, and VMB assessments, and adjustments for several known confounders. However, an even larger sample size and more frequent follow-up assessments would have been preferable. The study lacked quantitative HPV and VMB data at both time points, and HIV viral load at endline, and a full assessment of the HPV virome would have added value [26]. Some associations may have been detected due to chance because of the high dimensionality of our data and multiple testing. Finally, our results may not be generalisable to HIV-negative women, HIV-positive women on ART with sustained undetectable virus, and women in other world regions.

In conclusion, hrHPV exposure (and/or increased sexual risk-taking) may cause vaginal dysbiosis, but a bidirectional relationship is possible. At-risk women may therefore benefit from interventions that promote vaginal lactobacilli. In our population, vaginal dysbiosis does not increase CIN2+ risk, but CIN2+ increases vaginal dysbiosis risk, when hrHPV status is taken into account. Interventions that promote vaginal lactobacilli may therefore not prevent

CIN2+ development after a persistent hrHPV infection has taken hold. The potential association between progestin-only injectable contraceptive use and CIN2+ development deserves urgent attention.

Acknowledgements: This work was funded by the European Commission 7th Framework Programme [original HARP study; grant agreement number HEALTH-2010-F2-265396] and the University of Liverpool [16S rRNA sequencing]. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the authors' institutions or the funders. We thank the study participants, the HARP Study Group, and the Centre for Genomic Research and research support staff at the University of Liverpool. The HARP Study Group consisted of: A. Chikandiwa, E. Cutler, S. Delany-Moretlwe, L. Diep, C. Hwinya, D. A. Lewis, M.P. Magooa, V. Maseko, P. Michelow, B. Muzah, T. Omar, A. Puren (Johannesburg, South Africa); F. Djigma, J. Drabo, O. Goumbri-Lompo, D. Karambiri, N. Meda, B. Sawadogo, J. Simporé, R. Sombie, A. Yonli, Z. Zamane, S Zan (Ouagadougou, Burkina Faso); V. Costes, M.N. Didelot, S. Doutre, N. Leventoux, N. Nagot, J. Ngou, M. Segondy (Montpellier, France); and A. Devine, C. Gilham, L. Gibson, H. Kelly, R. Legood, P. Mayaud, H.A. Weiss (London, UK).

Author contributions: JvdW obtained the additional funding for the sub-studies. AG, AD, SF, PM, and JvdW wrote the study protocol. HK, AC, SDM, MS, TO, and PM collected the primary HARP data. AG and AD conducted the 16S rRNA sequencing and analysis. AG, MV, AD, and JvdW analysed the data. JvdW wrote the manuscript. All authors commented on and approved the final manuscript.

Declaration of interests: None of the authors report a conflict of interest.

References

1 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**:394–424.

2 Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, *et al.* Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999; **189**:12–19.

3 Bruni L, Diaz M, Castellsagué X, Ferrer E, Bosch FX, de Sanjosé S. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *J Infect Dis* 2010; **202**:1789–1799.

4 McDonald AC, Tergas AI, Kuhn L, Denny L, Wright TC. Distribution of human papillomavirus genotypes among HIV-positive and HIV-negative women in Cape Town, South Africa. *Front Oncol* 2014; **4**:48.

5 Ahdieh L, Klein RS, Burk R, Cu-Uvin S, Schuman P, Duerr A, *et al.* Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV-negative women. *J Infect Dis* 2001; **184**:682–690.

6 van de Wijgert JHHM, Jespers V. The global health impact of vaginal dysbiosis. *Res Microbiol* 2017; **168**:859–864.

7 Brusselaers N, Shrestha S, van de Wijgert J, Verstraelen H. Vaginal dysbiosis and the risk of human papillomavirus and cervical cancer: systematic review and meta-analysis. *Am J Obstet Gynecol* 2019; **221**:9-18.e8.

8 Kelly HA, Ngou J, Chikandiwa A, Sawadogo B, Gilham C, Omar T, *et al.* Associations of Human Papillomavirus (HPV) genotypes with high-grade cervical neoplasia (CIN2+) in a cohort of women living with HIV in Burkina Faso and South Africa. *PloS One* 2017; **12**:e0174117.

9 Kelly HA, Sawadogo B, Chikandiwa A, Segondy M, Gilham C, Lompo O, *et al.* Epidemiology of high-risk human papillomavirus and cervical lesions in African women living with HIV/AIDS: effect of anti-retroviral therapy. *AIDS* 2017; **31**:273–285.

10 *Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach: 2010 revision*. Geneva: World Health Organization; 2010. http://www.ncbi.nlm.nih.gov/books/NBK138540/ (accessed 8 May 2019).

11 *Guidelines: national consolidated guidelines for PMTCT and the management of HIV in children, adolescents and adults*. Department of Health RoSA; 2015. https://www.health-e.org.za/2015/07/02/guidelines-national-consolidated-guidelines-for-pmtct-and-the-management-of-hiv-in-children-adolescents-and-adults/ (accessed 8 May 2019).

12 Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, *et al.* A review of human carcinogens--Part B: biological agents. *Lancet Oncol* 2009; **10**:321–322.

13 Clement P, Stall J, Young R. In: *Atlas of gynecologic surgical pathology*.London, New York, Philadelphia, St Louis, Sydney, Toronto: Elsevier; 2014. p. Chapter 4.

14 Doutre S, Omar T, Goumbri-Lompo O, Kelly H, Clavero O, Zan S, *et al.* Cervical intraepithelial neoplasia (CIN) in African women living with HIV: role and effect of rigorous histopathological review by a panel of pathologists in the HARP study endpoint determination. *J Clin Pathol* 2018; **71**:40–45.

15 Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J Clin Microbiol* 1991; **29**:297–301.

16 Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* 2014; **2**. doi:10.7717/peerj.593

17 Jespers V, Crucitti T, Menten J, Verhelst R, Mwaura M, Mandaliya K, *et al.* Prevalence and correlates of bacterial vaginosis in different sub-populations of women in sub-Saharan Africa: a cross-sectional study. *PloS One* 2014; **9**:e109670.

18 van de Wijgert JHHM, Verwijs MC, Turner AN, Morrison CS. Hormonal contraception decreases bacterial vaginosis but oral contraception may increase candidiasis: implications for HIV transmission. *AIDS* 2013; **27**:2141–2153.

19 Marrinan JE, Chikandiwa A, Kelly HA, et al. The influence of hormonal contraception on high-risk HPV and high-grade cervical lesions among women living with HIV-1 in South Africa [abstract HPV17-0420]. In: 31st International Papillomavirus Conference, Cape Town, South Africa. http://hpv2017.org/scientific-information/abstracts-pdf#.XNUwHI4zZPa (accessed May 10, 2019).

20 Brotman RM, Shardell MD, Gajer P, Tracy JK, Zenilman JM, Ravel J, *et al.* Interplay between the temporal dynamics of the vaginal microbiota and human papillomavirus detection. *J Infect Dis* 2014; **210**:1723–1733.

21 Reimers LL, Mehta SD, Massad LS, Burk RD, Xie X, Ravel J, *et al.* The cervicovaginal microbiota and its associations with human papillomavirus detection in HIV-infected and HIV-uninfected women. *J Infect Dis* 2016; **214**:1361–1369.

22 Di Paola M, Sani C, Clemente AM, Iossa A, Perissi E, Castronovo G, *et al.* Characterization of cervico-vaginal microbiota in women developing persistent high-risk Human Papillomavirus infection. *Sci Rep* 2017; **7**:10200.

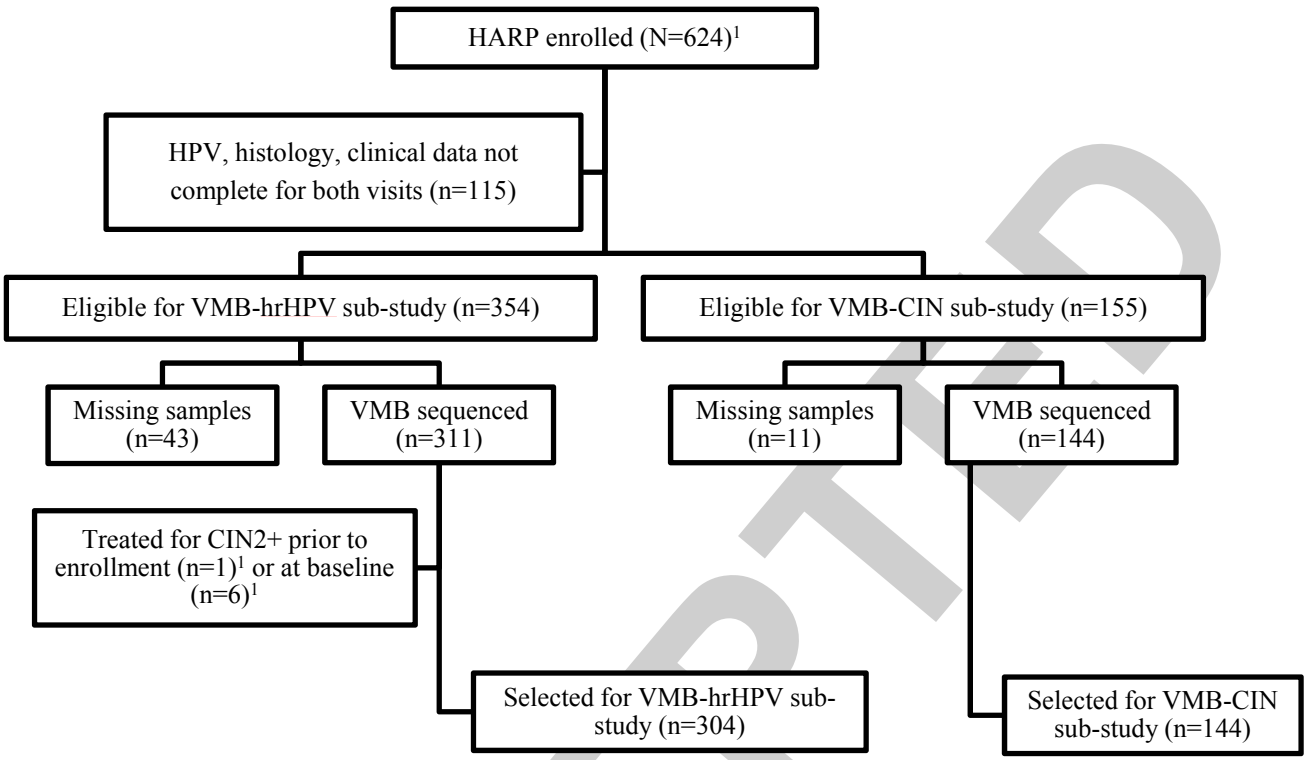
23 Karim R, Tummers B, Meyers C, Biryukov JL, Alam S, Backendorf C, *et al.* Human papillomavirus (HPV) upregulates the cellular deubiquitinase UCHL1 to suppress the keratinocyte's innate immune response. *PLoS Pathog* 2013; **9**:e1003384.

24 Ilhan ZE, Łaniewski P, Thomas N, Roe DJ, Chase DM, Herbst-Kralovetz MM. Deciphering the complex interplay between microbiota, HPV, inflammation and cancer through cervicovaginal metabolic profiling. *EBioMedicine* Published Online First: 24 April 2019. doi:10.1016/j.ebiom.2019.04.028

25 Borgdorff H, Gautam R, Armstrong SD, Xia D, Ndayisaba GF, van Teijlingen NH, *et al.* Cervicovaginal microbiome dysbiosis is associated with proteome changes related to alterations of the cervicovaginal mucosal barrier. *Mucosal Immunol* 2016; **9**:621–633.

26 Chen X, Kost J, Sulovari A, Wong N, Liang WS, Cao J, *et al.* A virome-wide clonal integration analysis platform for discovering cancer viral etiology. *Genome Res* 2019; **29**:819–830.

Figure 1: Study design and flow diagram

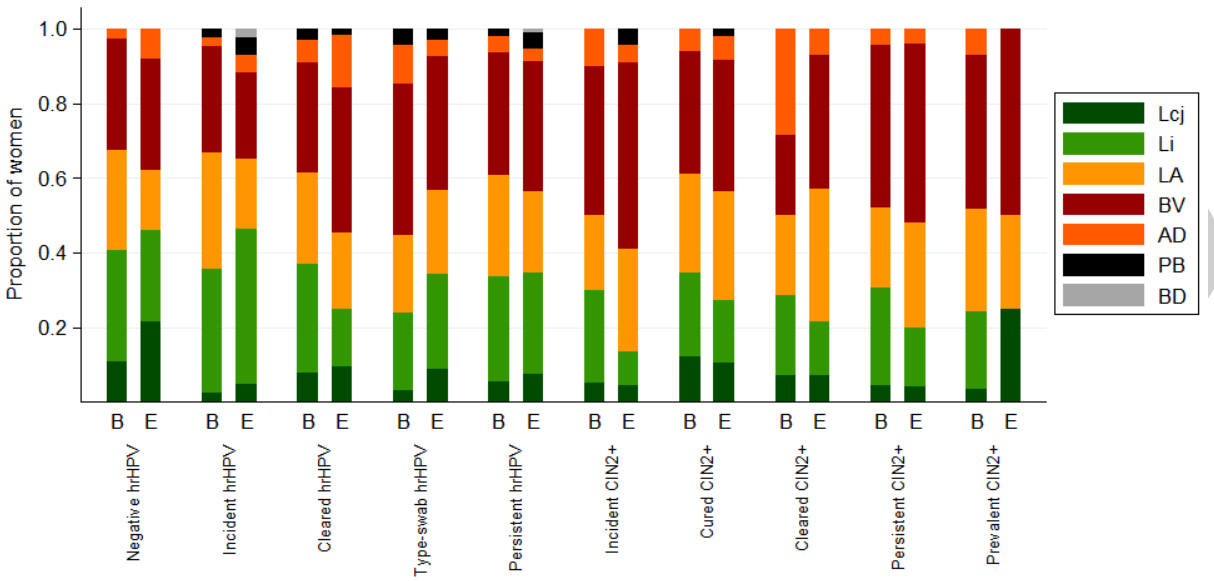


Enrolled in VMB sub-studies (N=448)			
VMB-hrHPV sub-study (N=304)	**n**	**VMB-CIN sub-study (N=144)**	**n**
≤CIN1 at both visits AND:		*Regardless of hrHPV status*	
Control group: No hrHPV at both visits	37	Control group: ≤CIN1 but same hrHPV at both visits[2]	92
Incident hrHPV	43	Incident CIN2+	22
Cleared hrHPV	65	Cured CIN2+ (after treatment)	50
Type swap	67	Cleared CIN2+ (spontaneously)	14
≤CIN1 but same hrHPV at both visits[2]	92	Persistent CIN2+	25
		[Prevalent CIN2+][3]	33

Abbreviations: *CIN* cervical intraepithelial neoplasia, *hrHPV* high-risk human papillomavirus, *VMB* vaginal microbiota. N and n refer to the number of women.

1. One woman had had cervical cancer treatment prior to enrolment and was erroneously enrolled into the HARP study (N=623 correctly enrolled women). In addition, six women were treated for CIN2+ at baseline based on local histology results but their baseline histology was subsequently downgraded to ≤CIN1 by the HARP Endpoint Committee. This became clear after the samples had already been sequenced. Data from these seven women were included in the bioinformatics pipeline but excluded from all epidemiological analyses.
2. These are the same 92 women. In this figure, they are included in the sample size of the VMB-hrHPV sub-study but not the sample size of the VMB-CIN sub-study. The total sample size of the VMB-CIN sub-study is therefore 144+92=236.
3. We required two full assessments per person except for women with CIN2+ at one visit and a missing histology result at the other visit. These are referred to as prevalent CIN2+ cases (n=29 at the baseline visit, and n=4 at the endline visit). They were included in some but not all analyses. Of the 59 women with missing endline data, 21 were lost to follow up, 16 moved away, ten were withdrawn by the clinician (including referral for suspected cervical carcinoma or hysterectomy), seven withdrew due to personal circumstances and/or being unwilling to undergo further testing and five died (two from HIV-related causes and three from other causes). The other women had incomplete HPV and/or CIN results.

Figure 2. VMB types at baseline and endline by hrHPV and CIN status



Abbreviations: *AD* BV-anaerobe-dominated, *B* baseline visit, *BD* *Bifidobacterium*-dominated, BV bacterial vaginosis-like, *CIN* cervical intraepithelial neoplasia, *E* endline visit, *hrHPV* high risk human papillomavirus, *LA* lactobacilli and BV-anaerobes, *Lcj Lactobacillus crispatus* or *L. jensenii*-dominated, *Li L. iners*-dominated, *PB* pathobionts-characterised, *VMB* vaginal microbiota.

The number of women is equal at the baseline and the endline visit in all outcome groups (see Figure 1 for sample sizes), with the exception of the prevalent CIN2+ groep: 29 women were seen at baseline only and four women were seen at endline only.

Table 1. Participant characteristics at baseline and endline (all groups combined)

Sociodemographic and behavioural characteristics	**Baseline (N=448)**	**Endline (N=432)**
Age (median, IQR)	34 (30-39)	NA
Current smoker (n %)	24 (5.4)	26 (6.0)
Has a current regular male sex partner (n %) – Cohabiting	362 (80.8) 182 (40.7)	NA NA
Number of male sex partners last 3 (baseline) or 6 (endline) months (n %) – 0 – 1 – 2+	 77 (17.2) 345 (77.0) 25 (5.6)	 55 (12.7) 359 (83.1) 18 (4.2)
Ever used hormonal contraception (n %)	386 (86.2)	369 (82.4)
Current use of hormonal contraception (n %)[1] – Combined pill (n %) – Combined patch (n %) – Progestin-only injectable (n %)	113 (25.2) 23 (5.1) 6 (1.3) 84 (18.8)	114 (27.1) 22 (5.2) 4 (1.0) 88 (20.9)
Any condom use in past two weeks (n %) – Always – Sometimes – Never – No sex	 218 (48.7) 125 (27.9) 24 (5.4) 81 (18.1)	 NA NA NA NA
Practices vaginal cleansing at least weekly (n %)	186 (41.5)	NA
Clinical and laboratory results		
Antiretroviral therapy: – On therapy at baseline and endline – Started therapy during the study – Not on therapy at all during the study		 300 (69.4) 23 (5.3) 109 (25.2)
CD4+ count in cells/μl (median, IQR)[2]	423 (317-566)	426 (324-565)
HIV undetectable in plasma (n %)[3]	103 (23.4)	NA
HIV plasma viral load $\log_{10}$ copies/ml (median, IQR)[3]	2.6 (1.6-3.9)	NA
Any high-risk human papillomavirus by genotyping	357 (79.7)	306 (70.8)
Yeasts on vaginal Gram stain (n %)	33 (7.5)	NA
Trichomonas vaginalis by NAAT (n %)	70 (15.6)	NA
Chlamydia trachomatis by NAAT (n %)	24 (5.4)	NA
Neisseria gonorrhoeae by NAAT (n %)	10 (2.2)	NA
Mycoplasma genitalium by NAAT (n %)	36 (8.0)	NA
Active syphilis by serology (n %)	3 (0.7)	NA
Herpes simplex virus type 2 by serology (n %)	425 (95.3)	NA
VMB characteristics		
Bacterial vaginosis by Nugent 7-10 (n %)	188 (42.9)	NA
Simpson diversity 1-D (mean, 95% CI)[4]	0.54 (0.51-0.57)	0.54 (0.51-0.57)
Lactobacilli relative abundance (mean, 95% CI)[4]	0.48 (0.44-0.52)	0.46 (0.42-0.50)
BV-anaerobes relative abundance (mean, 95% CI)[4]	0.49 (0.45-	0.50 (0.46-

	0.52)	0.54)
Pathobionts relative abundance (mean, 95% CI)[4]	0.03 (0.02-0.04)	0.03 (0.02-0.04)
Other bacteria relative abundance (mean, 95% CI)[4]	0.01 (0.01-0.01)	0.01 (0-0.02)
Vaginal microbiota type (n %):[4]		
- *Lactobacillus crispatus*/*L. jensenii*-dominated (Lcj)	27 (6.2)	38 (9.1)
- *L. iners*-dominated (Li)	115 (26.3)	95 (22.8)
- *Bifidobacterium*-dominated (BD)	0	2 (0.5)
- Lactobacilli + BV-anaerobes (LA)	111 (25.3)	95 (22.8)
- BV-like (BV)	148 (33.8)	149 (35.8)
- BV-anaerobe-dominated (AD)	29 (6.6)	26 (6.3)
- Pathobiont-characterised (PB)	8 (1.8)	11 (2.6)
- Did not align	0	1 (0.2)

Abbreviations: *BV* bacterial vaginosis, *CI* confidence interval, *IQR* inter-quartile range, *NA* not assessed, *NAAT* nucleic acid amplification test.

1.N=420 at endline.
2.N=381 at endline.
3.N=439 at baseline. If below the detection limit of 40 copies/ml, $\log_{10}$ of 40 was used.
4.N=438 at baseline and N=417 at endline.

Table 2: Bivariable multinomial logistic regression models with hrHPV and CIN outcomes

hrHPV outcome categories *All ≤ CIN1; control group: no hrHPV at both visits*	**Baseline VMB**	**RRR**	**p**	**Endline VMB**	**RRR**	**p**
Incident hrHPV	Lcj vs Li	0.196	0.171	Lcj vs Li BV vs Li	0.125 0.455	**0.019** 0.187
Cleared hrHPV	None			Simpson BV+AD+PB vs Li Lactobacilli RA BV-anaerobes RA	3.856 2.250 0.386 2.331	**0.034** 0.146 0.057 0.092
hrHPV type-swap	Nugent Simpson AD vs Li BV+AD+PB vs Li Lactobacilli RA BV-anaerobes RA	1.183 3.818 2.316 2.423 0.277 3.246	**0.016** **0.042** 0.136 0.090 **0.014** **0.026**	Lcj vs Li	0.397	0.174
Type-specific hrHPV persistence	BV-anaerobes RA	1.936	0.188	Simpson Lcj vs Li	2.716 0.315	0.184 0.074
Any hrHPV at one visit	None			Simpson Lcj vs Li	2.395 0.321	0.132 0.072
Any hrHPV at both visits	Nugent Simpson Lactobacilli RA BV-anaerobes RA	1.126 2.066 0.412 2.386	**0.050** 0.198 0.055 0.065	Simpson Lcj vs Li	2.165 0.348	0.162 0.069
CIN outcome categories *Control group: ≤CIN1 and hrHPV-positive at both visits*	**Baseline VMB**	**RRR**	**p**	**Endline VMB**	**RRR**	**p**
Incident CIN2+	None			Simpson LA vs Li BV vs Li BV+AD+PB vs Li Lactobacilli RA BV-anaerobes RA	7.352 3.751 4.297 4.167 0.335 2.828	**0.028** 0.129 0.073 0.075 0.086 0.101
Cured CIN2+	Lcj vs Li	2.836	0.139	LA vs Li	2.187	0.144
Cleared CIN2+	AD vs Li	8.662	**0.021**	Simpson LA vs Li	3.597 3.123	0.198 0.200
Persistent CIN2+	None			Simpson BV vs Li BV-anaerobes RA	4.239 2.343 2.291	0.068 0.181 0.149
CIN2+ at one or two visits	Nugent	1.087	0.170	Simpson	5.981	**0.003**
Control group: no hrHPV at both visits	AD vs Li Lactobacilli RA BV-anaerobes RA	4.259 0.433 2.561	0.187 0.075 **0.049**	LA vs Li BV vs Li BV+AD+PB vs Li Lactobacilli RA BV-anaerobes RA	3.094 2.404 2.210 0.352 2.412	0.064 0.101 0.122 **0.025** 0.060

Abbreviations: *BV* bacterial vaginosis-like, *CIN* cervical intraepithelial neoplasia, *hrHPV* high-risk human papillomavirus, *LA* lactobacilli and BV-anaerobes, Lcj *Lactobacillus crispatus* or *L. jensenii*-

dominated, Li *L. iners*-dominated, *RA* relative abundance, *RRR* relative risk ratio, *VMB* vaginal microbiota. Each model includes one multi-category outcome and one VMB independent variable. The VMB independent variable in each hrHPV outcome category is compared to that in the persistent hrHPV-negative control group. VMB variables tested: Nugent score (baseline only), Simpson diversity index, VMB type, and relative abundance of lactobacilli, BV-anaerobes, pathobionts, and other bacteria. Only results with $p<0.2$ are shown in the table: see Table S4 in Supplement 1 , http://links.lww.com/QAD/B541 for full results including 95% confidence intervals. 'None' means that none of the VMB variables were associated with the outcome listed in the first column at $p<0.2$. Statistically significant results ($p<0.05$) are shown in bold.

Table 3. Multinomial logistic regression models with combined hrHPV/CIN outcome

Bivariable models[1]	**Combined hrHPV/CIN outcome**[2]							
	All ≤CIN1; hrHPV at one visit vs no hrHPV at both visits		All ≤CIN1; hrHPV at both visits vs no hrHPV at both visits		CIN2+ at one or both visits vs ≤CIN1 and no hrHPV both visits		CIN2+ at one or both visits vs ≤CIN1 and hrHPV at both visits	
	RRR	**p**	**RRR**	**p**	**RRR**	**p**	**RRR**	**p**
Age (per year)	0.945	0.077	0.923	**0.010**	0.881	**<0.001**	0.954	**0.021**
Current smoker at baseline	1.385	0.774	2.939	0.308	1.839	0.574	0.626	0.339
Regular partner at baseline[3]	1.178	0.516	1.547	0.073	1.509	0.094	0.975	0.872
# sex partners 3 months prior to baseline[4]	1.877	0.113	1.773	0.131	1.919	0.090	1.082	0.752
On hormonal contraception at baseline	0.691	0.401	0.628	0.270	1.481	0.337	2.358	**0.001**
Any hormonal contraception during study[5]	1.074	0.671	0.923	0.627	1.216	0.229	1.317	**0.008**
Current condom use[6]	0.979	0.931	0.945	0.809	0.763	0.249	0.807	0.140
Vaginal cleansing at least weekly	0.997	0.987	0.820	0.295	0.878	0.493	1.070	0.580
On ART at baseline	1.877	0.138	1.179	0.674	0.800	0.568	0.678	0.114
CD4 count (per cell/µl)	1.000	0.946	0.999	0.348	0.998	0.053	0.999	0.128
Log_{10} HIV plasma viral load	0.830	0.257	1.026	0.869	1.231	0.175	1.200	**0.050**
CT, NG or syphilis[7]	3.673	0.223	1.658	0.641	3.876	0.198	2.338	0.076
Nugent score baseline[8]	1.049	0.444	1.126	**0.050**	1.087	0.170	0.965	0.367
Yeasts on Gram stain baseline[8]	2.059	0.511	3.525	0.231	3.256	0.265	0.924	0.847
Simpson index baseline	1.460	0.518	2.066	0.198	1.717	0.344	0.831	0.619
Simpson index endline	2.395	0.132	2.165	0.162	5.981	**0.003**	2.763	**0.012**
Lactobacilli RA baseline	0.593	0.276	0.412	0.055	0.433	0.075	1.051	0.865
Lactobacilli RA endline	0.589	0.252	0.651	0.333	0.352	**0.025**	0.541	**0.043**
BV-anaerobes RA baseline	1.725	0.266	2.386	0.065	2.561	**0.049**	1.073	0.809
BV-anaerobes RA endline	1.453	0.426	1.287	0.575	2.412	0.060	1.874	**0.039**
Pathobionts RA baseline	0.594	0.764	1.391	0.832	0.375	0.576	0.269	0.275
Pathobionts RA endline	452.02	0.275	524.55	0.262	376.63	0.289	0.718	0.744
Multivariable models[9]	All ≤CIN1; hrHPV at one visit vs no hrHPV at both visits		All ≤CIN1; hrHPV at both visits vs no hrHPV at both visits		CIN2+ at one or both visits vs ≤CIN1 and no hrHPV both visits		CIN2+ at one or both visits vs ≤CIN1 and hrHPV at both visits	
	aRRR	**p**	**aRRR**	**p**	**aRRR**	**p**	**aRRR**	**p**
Nugent score baseline[8]	1.042	0.518	1.114	0.081	1.073	0.262	0.963	0.348
Simpson baseline	1.381	0.585	1.941	0.246	1.653	0.391	0.852	0.671
Simpson endline	2.476	0.143	2.094	0.217	7.691	**0.002**	3.673	**0.004**
Lactobacilli RA baseline	0.587	0.267	0.404	0.051	0.417	0.067	1.034	0.911
Lactobacilli RA endline	0.534	0.207	0.631	0.339	0.325	**0.027**	0.515	**0.050**
BV-anaerobes RA baseline	1.729	0.264	2.414	0.063	2.738	**0.039**	1.134	0.677
BV-anaerobes RA endline	1.565	0.375	1.238	0.662	2.792	**0.045**	2.255	**0.017**

Abbreviations: *ART* antiretroviral therapy, *BV* bacterial vaginosis, *CIN* cervical intraepithelial neoplasia, *CT Chlamydia trachomatis*, *hrHPV* high-risk human papillomavirus, *NG Neisseria gonorrhoeae*, *RA* relative abundance, *aRRR* adjusted relative risk ratio. Statistically significant results ($p<0.05$) are shown in bold.

1. Each line represents one model. Each model includes the multi-category combined outcome and one independent variable.
2. With 'hrHPV' we mean detection of any type. With 'no hrHPV' we mean that not a single high risk type was present.
3. Comparing three categories: no, yes but not cohabiting, or cohabiting. The results are similar if

analysed as an indicator variable.

4. Comparing three categories: none, one, or two or more. The results are similar if analysed as an indicator variable.
5. Comparing three categories: never, at one visit, and at both visits. The results are similar if analysed as an indicator variable. When four categories are analysed as an indicator variable (never, baseline only, endline only, or at both visits), hormonal contraceptive use at baseline and at both visits are both associated with CIN2+ compared to hrHPV twice, but hormonal contraceptive use at endline only is not.
6. Comparing four categories: never, sometimes, always, or no sex in last six months. The results are similar if analysed as an indicator variable.
7. *Trichomonas vaginalis* was only associated with hrHPV twice (RRR=2.544, p=0.142, compared to negative controls). *Mycoplasma genitalium* was not associated with any outcome.
8. Not assessed at endline.
9. Each line represents one multivariable model. All multivariable models included the multi-category combined outcome, one VMB independent variable, age, CD4+ count at baseline or endline, and current hormonal contraceptive use (baseline) or any hormonal contraception during the study (endline). See Table S5 in Supplement 1, http://links.lww.com/QAD/B541 for 95% confidence intervals.