**Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Alexandra Vaia Ricca Date

# CHARACTERISTICS OF AN EVOLVING COHORT AND PREDICTORS OF HIV PREVALENCE AMONG HETEROSEXUAL DISCORDANT COUPLES IN LUSAKA, ZAMBIA

By

Alexandra Vaia Ricca
Master of Public Health

Global Epidemiology

_____

Susan Allen, MD MPH
Committee Chair

# CHARACTERISTICS OF AN EVOLVING COHORT AND PREDICTORS OF HIV PREVALENCE AMONG HETEROSEXUAL DISCORDANT COUPLES IN LUSAKA, ZAMBIA

By

Alexandra Vaia Ricca

Bachelor of Science
Miami University
2006

Thesis Committee Chair: Susan Allen, MD MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Epidemiology
2011

ABSTRACT

CHARACTERISTICS OF AN EVOLVING COHORT AND PREDICTORS OF HIV PREVALENCE AMONG HETEROSEXUAL DISCORDANT COUPLES IN LUSAKA, ZAMBIA

By Alexandra Vaia Ricca

Background: Many of the most important research findings on sexually transmitted infections (STIs) in the past 30 years have come from sub-Saharan Africa (1). Over 80% of all HIV-infected adults live in sub-Saharan Africa (2) and an estimated 340 million new cases of Chlamydia, gonorrhea and trichomoniasis are reported annually worldwide (3). There is a strong association of STIs, genital abnormalities and risky sexual behaviors with HIV acquisition.

Methods: Heterosexual discordant couples (one HIV-negative partner and one HIV-positive partner) from the Rwanda-Zambia HIV Research Group (RZHRG), recruited from voluntary counseling and testing (VCT) centers between 1994 and 2010, were used in secondary data analysis for descriptive statistics of data collected at baseline visit. The prevalence of gonorrhea, syphilis, trichomoniasis, urethral and vaginal discharge, HIV infection and other genital abnormalities and sexual behaviors were assessed. Data were restricted to 2890 couples (n=5780 individuals) with baseline information from past medical history, physical examination and laboratory testing. Couples were divided into three groups (enrolled in February 6, 1995 - January 22, 2002; March 15, 2002 - December 14, 2006; or January 5, 2007 - February 25, 2010) to compare prevalence of STIs over time and with different diagnostic procedures. Multivariate logistic regression was used as the primary method in determining the set of STI, genital abnormality and sexual behavior variables that most strongly predicted HIV prevalence at baseline.

Results: Across all groups, men had significantly more lifetime sexual partners (mean=11.15) than women (mean=3.40). Men reported more gonorrhea, ulceration and cystitis/dysuria in the past compared to women. Women reported more cases of syphilis, discharge and lower abdominal pain in the past as well as had inflammation, discharge and ulceration noted on a physical exam than men. Individuals with HIV infection were more likely to have had gonorrhea, syphilis, ulceration, discharge and lower abdominal pain in the past. Past reports of gonorrhea (OR: 2.46, 95% CI: 1.07, 5.66), ulceration (OR: 1.98, 95% CI: 1.42, 2.76), cystitis/dysuria (OR: 1.40, 95% CI: 1.01, 1.94) and a positive syphilis serology (OR: 1.77, 95% CI: 1.77, 95% CI: 1.38, 2.27) were significantly associated with an increased risk of HIV prevalence at baseline across all three groups, stratified by sex.

Discussion: Significant differences of STIs, genital abnormalities and sexual behavior exist between men and women and HIV negative and HIV positive individuals within study entry group. The presence of ulcerative and non-ulcerative STIs is associated with an increased risk of HIV infection.

# CHARACTERISTICS OF AN EVOLVING COHORT AND PREDICTORS OF HIV PREVALENCE AMONG HETEROSEXUAL DISCORDANT COUPLES IN LUSAKA, ZAMBIA

By

Alexandra Vaia Ricca

Bachelor of Science
Miami University
2006

Thesis Committee Chair: Susan Allen, MD MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Epidemiology
2011

**TABLE OF CONTENTS**

# LIST OF TABLES

# BACKGROUND

Many of the most important research findings in the past 30 years on sexually transmitted infections (STIs) have come from sub-Saharan Africa (1) where over 80% of all HIV-infected adults reside (2). An estimated 340 million new cases of chlamydia, gonorrhea, syphilis and trichomoniasis are reported annually throughout the world and almost 57 million disability-adjusted life years (DALYs) are lost to STIs (3). Unprotected sexual intercourse, multiple sex partners and other risky sexual behaviors place individuals at a high risk for acquisition of STIs and HIV.

A strong association exists between the spread of STIs and HIV transmission where the risk of sexual HIV transmission is increased with various ulcerative and non-ulcerative STIs (4) and observational studies have consistently shown the role of STIs in HIV transmission facilitation (5). Importantly, it has been identified that a higher relative risk of HIV acquisition is associated with genital ulcer disease, such as herpes simplex virus type 2 (HSV-2) compared to other STIs (6). Epidemiological synergy is a crucial component in STI research and many studies have shown the roles other infections play in HIV acquisition and affecting the natural history of the virus (7). Three relationships have been postulated to demonstrate HIV and STI co-infection: increased transmission and accelerated progression of HIV, alteration in the natural history, and diagnosis or response to therapy of HIV in the presence of other STIs (7). Several studies have shown the importance STIs have in predicting HIV acquisition. One such study from South Africa resulted in gonorrhea being associated with a positive HIV serostatus among all

men with urethral discharge syndromes (8). In the Rakai study, one out of the two most important risk factors found for HIV transmission was the presence of genital ulcers among HIV discordant, monogamous couples (6). Notably, higher transmission probabilities per coital act have been found among individuals who have reported genital ulceration (9). A study in 2003 among high-risk women in Lusaka, Zambia found that women who seroconverted from a baseline HIV-negative status during follow-up were significantly more likely to be infected with vaginal Candida (10). Other studies have confirmed significant associations between HIV seropositivity with Candida, bacterial vaginosis, infection with Trichomonas vaginalis, discharge/dysuria, genital ulcers and other disturbances of the normal vaginal flora (11-13). In a study among new HIV infections in Zimbabwean and Ugandan women, it was found that a substantial proportion of HIV seroconversion was attributable to reproductive tract infections, particularly HSV-2 and altered vaginal flora (5). Generally, it is understood that at the population level, controlling STIs contributes to a reduction in HIV incidence (5). Furthermore, given the evidence of synergy between HIV transmission and STIs, prevention methods targeting untreated STIs offer an opportunity for altering the natural history and course of the HIV epidemic in sub-Saharan Africa (14).

Voluntary counseling and testing (VCT) has illustrated cost-effectiveness and beneficial in some risk groups, particularly in cohabiting couples where most new HIV infections occur (15). Zambia's adult HIV prevalence rate is approximately 15.2 % (16) where many of the virus' transmission in urban Zambia takes place within marriage or cohabitation, with one in five cohabiting couples having one HIV-positive and one HIV-

negative partner (‘discordant couples’) (2). Diagnoses of STIs and genital abnormalities differ between countries, clinics and across follow up years of research. Syndromic management of STIs is common in research and clinical settings of sub-Saharan Africa, mainly due to the lack of resources, but have been shown to decrease prevalence of symptomatic STIs (17). The WHO guidelines note components of case management, defined as the “care of a person with an STI-related syndrome or with a positive test for one or more STIs,” including but not limited to, history taking, clinical examination, correct diagnosis, early and effective treatment, counseling and follow-up (4). When laboratory confirmation testing for STIs and HIV is available, physical examinations and laboratory tests are done.

Not only have STIs been found to promote HIV transmission, multiple recent sexual partnerships and risky sexual practices have been found to increase HIV acquisition as well. It appears that multiple sex partners is related to unprotected sex and the rapid spread of HIV in southern Africa has in part been attributed to multiple sex partners (18). Concurrent, or overlapping sex partners, also increases this risk in acquiring HIV and other STIs and by having multiple sex partners, risk of exposure to new HIV infection is increased.

The Zambia-Emory HIV Research Project (ZEHRP) opened in 1994 in Lusaka, Zambia as a second site to the Rwanda Zambia HIV Research Group (RZHRG), founded by Dr. Susan Allen. Dr. Allen began her work in HIV in Kigali, Rwanda in 1986, and has since promoted couples voluntary counseling and testing, methods of prevention for mother-to-

child transmission and anti-retroviral therapy programs. The sero-discordant couple cohorts from Rwanda and Zambia are the largest in the world (19).

## METHODS

The purpose of this study was to describe prevalence of STIs and genital abnormalities as well as sexual behaviors that have been exposed to changes in reporting and diagnostic methods in an evolving cohort that has been established for 15 years. We hypothesized that differences exist between men and women and between HIV-negative and HIV-positive individuals in each of three different reporting periods as well as across all reporting periods. We also hypothesized that ulcerative and non-ulcerative STIs as well as risky sexual behaviors and multiple sex partners are predictive of HIV prevalence at baseline.

According to RZHRG's research initiatives, the study on heterosexual transmission of HIV in Africa is an observational study to understand the behavioral, virologic, immunologic, and immunogenetic correlates of HIV transmission in HIV discordant couples. Laboratory testing and data collection is also done for other STIs, such as syphilis, gonorrhea, HSV-2, Candida, bacterial vaginosis and trichomoniasis. Sexually transmitted infections and genital abnormalities in low resource settings are significant health indicators in population health. RZHRG physicians and nurses have maintained medical information of their patients which include important data on sexual and reproductive health, including history of sexual activity and the number of partners in the past, condom use, baseline HIV status, vaginal and urethral discharge, external and internal ulceration and inflammation reported in the past and found in physical examinations, gonorrhea, and syphilis.

**Ethical considerations**

Emory University's Institutional Review Board (IRB) and the Zambian Ethics Committee approved the heterosexual transmission study in which this analysis took place. Approval of this analysis was provided by the RZHRG offices and was considered to be covered under the pre-existing RZHRG protocols. Informed consent was obtained by RZHRG at the time of enrollment for all study participants at ZEHRP in Lusaka, Zambia.

**Study area**

Zambia is a land-locked country in southern African with a population of approximately 14 million people (16). The country ranks among the highest in the world for HIV/AIDS and the highest disease burden can be found in the capital city of Lusaka where 45% of childbearing women reside and one in four adults have HIV infection (2). According to 2009 estimates for Zambia, approximately 980,000 individuals are living with HIV/AIDS (16).

**Study population**

Beginning in 1994, couples were recruited at VCT centers around Lusaka by the recruitment team where they watch a video, participate in a discussion session and are counseled with a trained counselor prior to HIV testing (15). Study materials, discussion sessions and counseling were conducted in Nyanja, the local language. Couples giving written consent for joint testing had their blood drawn and were assigned an identification number that was linked to their test results (15). Both the man and the woman were tested using rapid HIV tests and rapid plasma reagin (RPR) serology on-site, followed by joint

posttest counseling (15). Those couples with concordant results were invited to return for repeat tests or follow-up counseling at any time (15). Discordant couples were invited to join ZEHRP's study.

Discordant couples interested in joining the study were asked to attend an enrollment visit to complete medical questionnaires and have their blood drawn. Data were collected separately for men and women by counselors of the same sex to insure confidentiality and accurate reporting (15, 20). Reimbursement of travel costs, monetary incentives, child care and lunch were provided at enrollment and follow-up visits. Essential drugs such as antibiotics, antimalarials and anti-inflammatories were stocked at the on-site pharmacy. Enrollment visits were offered Tuesdays through Saturdays and have remained constant over time. The mean duration of follow-up was 15 months and study benefits continued for all discordant couples as long as quarterly appointments were followed (15).

This analysis included data from 2890 couples (N=5780 individuals) in the Zambian heterosexual HIV transmission study cohort consisting of HIV discordant cohabiting couples in a sexual relationship for at least 6 months and residence in Lusaka at the time of enrollment (19). Age inclusion criteria included women between 16 and 45 years of age and men between 16 and 65 years of age. HIV status was determined with an algorithm involving two rapid antibody tests (2). Partners jointly consented to the study and confidentiality was secured through assignment of one study number per couple (HTID) (2). Each couple was provided with a sexual diary and explained how to record

sexual intercourse with and without condoms (2). Data for the current analysis came from past medical history reports, physical examinations and laboratory results recorded for the couples' baseline visit.

**Data collection**

All medical forms were completed by trained medical officers who interviewed study participants of the same sex (20). Data was entered on-site into Microsoft Access databases by trained data analysts. Past medical history, physical examination and laboratory tests were completed at the study participants' baseline visit.

*Past medical history*

Past medical history prior to 2002 included data for men and women on incidences of venereal disease in the 5 years prior to study visit. If they reported having venereal disease in the past 5 years, participants specified cases of gonorrhea, syphilis or ulceration. Men were asked to report urethral discharge, cystitis/dysuria, and lower abdominal pain in the past year. Women were asked to report vaginal discharge, cystitis/dysuria or lower abdominal pain cases in the past year. Past medical history forms after 2002 included data for men and women on urethral (for men) and vaginal (for women) discharge, cystitis/dysuria, lower abdominal pain, gonorrhea, syphilis, and acute and chronic ulceration during the year prior to the baseline study visit.

*Physical examination*

The physical examination form collected information from men through genital examinations and women through gynecologic examinations. During all three time periods, data on external genitalia inflammation, external genitalia ulceration,

inflammation of the cervix and vagina, ulceration of the cervix and vagina, and discharge of the cervix, vagina and urethra were collected. Composite variables were included in this dataset for analysis purposes. The ulceration composite variable was composed of ulcer variables or variables indicative of ulcer on the physical examination form (for men, this variable included external genital ulceration and for women, this variable included external ulceration, ulceration of the cervix, ulceration of the vagina or erosion/friability of the cervix or vagina). The inflammation composite variable was composed of inflammation variables or variables indicative of inflammation on the physical examination form (for men, this variable included inflammation or urethral discharge and for women, this variable included external inflammation, inflammation of the cervix or vagina or discharge of the cervix or vagina).

*Laboratory*

The serum database prior to 2002 included data on Trichomonas vaginalis wet prep, syphilis test results by Rapid Plasma Reagin (RPR) with confirmation by Treponema pallidum hemagglutination (TPHA) (2), gonorrhea test results (gram stain and culture), Candida by wet prep and bacterial vaginosis by wet prep. The database between 2002 and 2006 included data on syphilis test results (RPR), HSV-2 tests results, and wet prep for Trichamonas vaginalis. The database post-2007 included data on RPR, TPHA, HSV-2 and wet prep for Trichomonas vaginalis. For data analysis purposes, syphilis results by RPR only were used. Gonorrhea test data included results from gram stain and culture. Candida test data included results from wet prep and gram stain. TPHA confirmation was not done systematically post-2002, as treatment was provided regardless of RPR titer.

**Data analysis**

The primary objective was to consider frequencies of STI-related variables in past medical history, physical examinations and laboratory in Zambia during three different time periods between February 6, 1995 and January 22, 2002 (HTID 1-1634, group 1), March 15, 2002 and December 14, 2006 (HTID 1636-2994, group 2) and between January 5, 2007 and February 25, 2010 (HTID 2995-3976, group 3) to provide a snapshot of the cohort. Descriptive statistics between men and women and HIV-negative and HIV-positive individuals were assessed. The secondary objective was to build a model to estimate HIV prevalence at enrollment based on presence of STIs, genital abnormalities and sexual behaviors.

This study was performed as secondary analysis. The dataset of use was compiled in the Rwanda-Zambia HIV Research Group office in the School of Medicine, Department of Pathology at Emory University in Atlanta, Georgia. Variables of interest from past medical history forms, physical examination forms and the serum database were identified and the dataset was created with the 1994 to 2010 Zambian heterosexual transmission study cohort. Data were analyzed using SAS software package (version 9.2, Cary, NC).

For data analysis, dichotomous STI and genital abnormality variables for gonorrhea, syphilis, ulceration, discharge, cystitis/dysuria and lower abdominal pain reported in the past 5 years or past one year, noted on the physical examination form, or diagnosed by laboratory testing were created. Ulceration variables (cervical, vaginal and external

genitalia) were grouped together as one dichotomous ulceration variable as were inflammation variables (cervical, vaginal and external genitalia) from physical examinations. Vaginal discharge in women and urethral discharge in men were also grouped together in to one dichotomous variable for discharge. Composite variables in the original dataset from physical examinations that suggested inflammation and ulceration were also included in the analysis. Multiple laboratory tests for the same infection were grouped together into a dichotomous variable. Gonorrhea by gram stain or culture were made into one variable for positive or negative gonorrhea laboratory test. A variable for HSV-2 was compiled from two opportunities for reporting a positive HSV-2 result in the serum database. Candida wet prep and Candida by gram stain were included in one laboratory test for Candida.

Descriptive statistics and frequencies were found between men and women and HIV-negative and HIV-positive individuals for genital inflammation (including external for men and women and cervical and vaginal for women), ulceration (including external for men and women and cervical and vaginal for women), discharge (including urethral for men and vaginal and cervical for women), cystitis/dysuria, lower abdominal pain, diagnoses by laboratory tests of gonorrhea, syphilis, HSV-2, Candida, bacterial vaginosis for women, Trichomonas, partnerships outside of the main project partner and condom use. A total of 5780 individuals were included in the analysis, 2890 men and 2890 women.

The primary outcome in this study was prevalent HIV infection at baseline. This outcome was predicted by the existence of STI symptoms and ulcerative and non-ulcerative STIs as well as past sexual behavior, stratified by sex. Each STI, genital abnormality and sexual behavior variable was first entered into univariate logistic regression to evaluate crude associations with HIV prevalence at baseline. Odds ratios (OR) that were statistically significant ($p<0.05$) were entered into the multivariate model. Multivariate logistic regression was used as the primary method in determining the set of STI, genital abnormality and sexual behaviors variables that most strongly predicted HIV prevalence at baseline. Collinearity was diagnosed and variables with condition indices higher than 30 were considered to have a collinearity problem. Stepwise logistic regression was then used to eliminate variables that were not statistically significant in the full, adjusted model; 95% confidence intervals (CI) were calculated for the ORs.

# RESULTS

An equal number of HIV-positive and HIV-negative individuals were enrolled in the Zambian study based on the design of the discordant heterosexual transmission study. Table 1 describes basic demographic and behavior characteristics including age, number of partners and condom use. Zambian men were significantly older ($p<0.01$) in each of the three groups (group 1 mean=34.09, group 2 mean=35.39, group 3 mean=35.77) than their wives (group 1 mean=27.03, group 2 mean=29.03, group 3 mean=29.46). Zambian men were significantly older ($p<0.01$) at age of first sex in each of the three groups (group 1 mean=17.64, group 2 mean=17.49, group 3 mean=17.68) than their wives (group 1 mean=16.27, group 2 mean=16.62, group 3 mean=16.90). Zambian men had significantly more ($p<0.01$) lifetime sexual partners (group 1 mean=9.59, group 2 mean=13.97, group 3 mean=9.45) than their wives (group 1 mean=3.10, group 2 mean=3.66, group 3 mean=3.47). Zambian men have had significantly more ($p<0.01$) sexual partners in the past year (group 1 mean=1.80, group 2 mean=1.74, group 3 mean=1.52) than their wives (group 1 mean=1.12, group 2 mean=1.11, group 3 mean=1.18). Zambian men have also had significantly more ($p<0.01$) sexual partners in the past 3 months (group 1 mean=1.18, group 2 mean=1.08, group 3 mean=1.07) than their wives (group 1 mean=1.01, group 2 mean=0.97, group 3 mean=1.00). Table 2 describes demographic, behavior and health characteristics between HIV-positive and HIV-negative individuals. The couple HIV status at baseline differed across all groups. There were 1134 (52.79%) HIV-positive men with HIV-negative female partners in group 1, 1016 (45.64%) in group 2 and 566 (40.26%) in group 3. There were 1014

(47.21%) HIV-negative men with HIV-positive female partners in group 1, 1210 (54.36%) in group 2 and 840 (59.74%) in group 3. Generally, individuals who were HIV-positive and HIV-negative were the same age, had sexual debut at the same age, had the same number of lifetime sexual partners, partners in the past year and partners in the past 3 months. There were no statistically significant differences found between HIV-negative and HIV-positive individuals who had sexual intercourse with another partner, had sexual intercourse with another partner with or without a condom or condom use with their main project partner.

**Past medical history**

Previous reports of STIs differed between men and women in all three groups (Table 3). Differences were statistically significant ($p<0.05$) with the exception of syphilis in group 3. Generally, men more frequently reported past cases of gonorrhea, ulceration and cystits/dysuria whereas women more frequently reported past cases of syphilis, discharge and lower abdominal pain. Reporting past ulceration in men was highest (25.07%) whereas reporting past gonorrhea in women was lowest (0.28%).

Previous reports of STIs differed among individuals who were HIV-negative and HIV-positive in all three groups (Table 4). Statistically significant differences ($p<0.05$) were found in past gonorrhea cases in group 1 and across all groups, past syphilis cases in groups 1, 2 and across all groups, past ulceration cases in groups 1, 2, 3 and across all groups, past discharge cases in group 2 and across all groups, past cystitis/dysuria cases in groups 1, 2 and across all groups and finally past cases of lower abdominal pain in

group 1 and across all groups. Syphilis was reported the most in group 2 among HIV-negative individuals (15.81%) and ulceration in group 2 among HIV-positive individuals (25.61%). Lower abdominal pain was reported the least in group 3 among HIV-negative individuals (0.28%) and among HIV-positive individuals (0.71%).

Reports of past gonorrhea (OR: 1.35, 95% CI: 1.01, 1.79), syphilis (OR: 1.84, 95% CI: 1.43, 2.38), ulceration (OR: 2.14, 95% CI: 1.64, 2.78), cystitis/dysuria (OR: 1.74, 95% CI: 1.37, 2.20) or lower abdominal pain (OR: 1.58, 95% CI: 1.26, 1.97) in the first group were significantly associated with HIV prevalence at baseline (Table 5). Syphilis (OR:1.73, 95% CI: 1.40, 2.13), ulceration (OR: 3.06, 95% CI: 2.42, 3.87), discharge (OR: 1.49, 95% CI: 1.07, 2.10) or cystitis/dysuria (OR: 1.48, 95% CI: 1.09, 2.00) significantly predicted HIV prevalence at baseline in group 2. In group 3, ulceration in the past was strongly associated with HIV prevalence at baseline (OR: 3.63, 95% CI: 2.24, 5.88).

**Physical examination**

Reported findings of genital abnormalities and symptoms associated with STIs differed between men and women in all three groups (Table 3). Differences were statistically significant ($p<0.05$) with the exception of inflammation in group 3, ulceration in groups 1, 2 and across all groups, discharge in group 3, symptoms suggestive of inflammation (inflammation composite variable) in groups 2 and 3 and symptoms suggestive of ulceration (ulceration composite variable) in group 2. In general, women were found to have more cases of inflammation, discharge, symptoms suggestive of inflammation (inflammation composite variable) and symptoms suggestive of ulceration (ulceration

composite variable) as opposed to men having more cases of ulceration. Symptoms suggestive of inflammation (inflammation composite variable) among women was most common (19.37%) and discharge among women was the least common (0.43%).

Reported findings of genital abnormalities and symptoms associated with STIs differed between HIV-negative and HIV-positive individuals in all three groups (Table 4). Differences were statistically significant ($p<0.05$) with the exception of inflammation in group 1 and 3, discharge in group 1 and 3, and symptoms suggestive of inflammation (inflammation composite variable) in groups 1, 2, 3 and across all groups. Inflammation in the composite variable was reported the most in group 1 among HIV-negative individuals (11.82%) and among HIV-positive individuals (12.85%). Interestingly, inflammation was reported the least in group 2 among HIV-negative individuals (0.27%) and inflammation in the composite variable in group 3 among HIV-positive individuals (0.43%).

Inflammation (OR: 1.55, 95% CI: 1.04, 2.31), ulceration (OR: 2.93, 95% CI: 1.96, 4.38) and symptoms suggestive of ulceration (ulceration composite variable) (OR: 2.19, 95% CI: 1.60, 2.99) in group 1 were strongly associated with HIV prevalence at baseline in univariate odds ratio analyses (Table 5). These predictors were also significant in predicting HIV prevalence at baseline in group 2 and across all groups. Ulceration (OR: 8.76, 95% CI: 2.56, 29.95) and the ulceration composite variable (OR: 4.95, 95% CI: 1.82, 13.47) were significantly associated with HIV prevalence at baseline in group 3 in univariate odds ratio analyses.

**Laboratory**

Laboratory results in the serum database differed among men and women in all groups (Table 3). Women were more likely to have a positive test for gonorrhea, syphilis, HSV-2, Candida and trichomoniasis. Approximately 19% of women in group 1 had a positive test for bacterial vaginosis. The highest percentage of positive syphilis by RPR was women (24.58%) in group 1 and the lowest percentage of positive gonorrhea, Candida and trichomoniasis tests were found in men among all groups (0.03%).

Laboratory results in the serum database differed between HIV-negative and HIV-positive individuals in all three groups (Table 4). Differences were statistically significant ($p<0.05$) with the exception of syphilis in group 3, HSV-2 in group 3, Candida in group 1 and across all groups, bacterial vaginosis in women in group 1 and across all groups and trichomoniasis in groups 2, 3 and across all groups. Syphilis was reported the most in group 1 among HIV-negative individuals (16.76%) and among HIV-positive individuals (25.23%). HSV-2 was reported the least in group 2 among HIV-negative individuals (0.09%) and in group 3 among HIV-positive individuals (0.14%).

Syphilis was significantly associated ($p<0.05$) with HIV prevalence at baseline in group 1 (OR: 1.72, 95% CI: 1.39, 2.13) as well as a positive test for trichomoniasis in group 1 (OR: 1.51, 95% CI: 1.11, 2.06) (Table 5). In group 2, positive syphilis (OR: 2.04, 95% CI: 1.64, 2.54) and HSV-2 (OR: 21.85, 95% CI: 1.69, 282.42) were associated with an increased risk for HIV prevalence at baseline. Across all groups, positive syphilis tests (OR: 1.72, 95% CI: 1.50, 1.98) and positive tests for trichomoniasis (OR: 1.26, 95% CI:

1.01, 1.57) were significantly associated with predicting HIV prevalence at baseline in univariate odds ratio analyses.

**Sexual behavior**

Sexual behavior variables differed significantly between men and women across all groups (Table 1). In general, men were significantly ($p<0.01$) more likely to report having sex outside of their project partner than women in the 3 months prior to baseline visit. They were also significantly ($p<0.01$) more likely to report not using a condom with an outside partner in the past 3 months. Interestingly, a significantly greater number of men reported condom use among sex with an outside partner in the past 3 months than their female partners. Generally, no significant difference was found between men and women and their condom use with their project partners in the 3 months prior to baseline visit. Sexual behavior variables differed slightly between HIV-negative and HIV-positive individuals across all groups (Table 2). A significant difference ($p<0.05$) was found in the mean age between HIV-negative (33.20) and HIV-positive (32.05) in group 3, the mean age of first sex between HIV-negative (17.51) and HIV-positive (17.07) in groups 3, mean number of lifetime partners between HIV-negative (5.88) and HIV-positive (6.78) in group 1 and the number of partners in the past 3 months between HIV-negative (1.03) and HIV-positive (1.07) individuals across all groups.

Sexual intercourse with a partner other than the main project partner in the 3 months prior to baseline study visit in group 3 (OR: 1.82, 95% CI: 1.18, 2.90) and across all groups

(OR: 1.50, 95% CI: 1.12, 2.02) was significantly associated ($p<0.01$) in predicting HIV prevalence at baseline in univariate odds ratio analyses (Table 5).

**Multivariate analysis**

After adjusting for other variables in the group 1 model, reporting gonorrhea in the past was associated with a 1.06 to 6.65-fold increase in risk for HIV prevalence at baseline (Table 6). Reporting cystitis/dysuria in the past was associated with a 1.02 to 2.38-fold increase in risk for HIV prevalence at baseline and reporting lower abdominal pain in the past was associated with a 1.09 to 2.24-fold increase in risk for HIV prevalence at baseline. In group 2, having a positive HSV-2 result from the laboratory was significantly associated with an increase in risk for HIV prevalence at baseline. A patient reporting ulceration in the past had a 1.84 to 5.08-fold increase in risk for HIV prevalence at baseline in group 3 and a 1.35 to 17-fold increase in risk for HIV prevalence at baseline in individuals having ulceration in the physical exam. The final model adjusting for other variables across all groups stratified by sex yields an increased risk for HIV prevalence at baseline with past gonorrhea (OR: 2.46, 95% CI: 1.07 to 5.66), past ulceration (OR: 1.98, 95% CI: 1.42 to 2.76), past cystitis/dysuria (OR: 1.40, 95% CI: 1.01 to 1.94) and positive syphilis laboratory test (OR: 1.77, 95% CI: 1.38 to 2.27).

# DISCUSSION

This study confirmed the hypothesis that differences exist between men and women and between HIV-negative and HIV-positive individuals within each specified reporting time group respective to entry into study as well as across groups. It also confirmed that ulcerative and non-ulcerative STIs and genital abnormalities are predictive of HIV prevalence at baseline, separately and in the presence of other predictors. This study did not confirm the hypothesis that risky sexual behaviors and multiple sex partners predict HIV prevalence at baseline.

The analysis found significant associations from past medical reports, physical examinations and laboratory results within each group and across all groups with HIV prevalence at baseline. In group 1, reporting past gonorrhea, syphilis, ulceration, cystitis/dysuria, lower abdominal pain, finding through examination inflammation, ulceration, symptoms suggesting ulceration and positive syphilis and trichomoniasis tests significantly predicting HIV prevalence at baseline. Group 2 yielded significant associations with HIV prevalence at baseline in past medical reports of syphilis, ulceration, discharge and cystitis/dysuria, physical examinations of inflammation, ulceration and symptoms suggesting ulceration as well as a positive syphilis and HSV-2 result. Reporting ulceration in the past and on a physical exam as well as symptoms suggestive of ulceration, along with having sex outside of a project partner was significantly associated with HIV prevalence at baseline in group 3. Significant predictors of HIV prevalence at baseline were found when all groups were combined, including

reporting gonorrhea, syphilis, ulceration, cystitis/dysuria and lower abdominal pain in the past. Physical examination reports of inflammation, ulceration and symptoms suggestive of ulceration and positive syphilis and Trichomonas laboratory results were also significant in predicting baseline HIV prevalence.

Differences can be partially explained by the lack of data in some groups due to discontinuation of a variable or a laboratory test. For example, gonorrhea, Candida and bacterial vaginosis laboratory tests were discontinued post-2002, so are only considered in group 1. Data on HSV-2 was insufficient in groups 1 and 2. Having sex with another partner was only asked during time periods 2 and 3 and condom use during sex with another partner was only asked in time period 1.

Interestingly, significant differences were not consistently found between HIV-negative and HIV-positive individuals by sexual behavior. Perhaps these results are a reflection of the nature of this study and enrolling only discordant HIV couples, confirming the uniqueness of this cohort and the need for continuing attention and promotion of couples voluntary counseling and testing programs. These questions are asked at baseline and are reflective of sexual activity prior to study entry, before couples are exposed to condom promotion and education.

There are limitations to this study. The lack of data in some time periods and discontinuation of certain laboratory tests or questions asked of sexual history present a significant gap in this research. Comprehensive STI diagnostic methods and data

collection are important in the control of Zambians health and transmission of HIV. Although not statistically significant, some sexual behavior was associated in predicting HIV prevalence at baseline, suggesting the need for improved access to condoms and education on preventing transmission of HIV between discordant couples. Overwhelmingly, couples in this cohort reported not using condoms with or without their project partner, suggesting a gap in education and behavior. However, because these questions were reported for the time prior to baseline visit, further investigation of condom use trends during follow-up should be assessed.

In conclusion, this study has highlighted the importance of STIs, genital abnormalities and sexual behavior in HIV acquisition. It has been demonstrated that the presence of some STIs and genital abnormalities are significantly associated with HIV acquisition prior to enrolling in a research study, confirming the necessity of STI and genital abnormality control and prevention efforts. From the public health perspective, it is important that RZHRG continues its success in recruiting and following discordant heterosexual couples in Lusaka, Zambia. Moving forward, the organization should consider adding more laboratory tests for common STIs and genital abnormalities and improving the method of data collection and management of the laboratory database to increase the amount of data available to determine appropriate interventions.

## REFERENCES


1. Mabey D, Ndowa F, Latif A. What have we learned from sexually transmitted infection research in sub-Saharan Africa? *Sex Transm Infect* 2010;86(7):488-92.
2. Allen S, Meinzen-Derr J, Kautzman M, et al. Sexual behavior of HIV discordant couples after HIV counseling and testing. *AIDS* 2003;17(5):733-40.
3. Ng BE, Butler LM, Horvath T, et al. Population-based biomedical sexually transmitted infection control interventions for reducing HIV infection. *Cochrane Database Syst Rev* 2011;3:CD001220.
4. Guidelines for the management of sexually transmitted infections. Geneva: World Health Organization, 2003.
5. van de Wijgert JH, Morrison CS, Brown J, et al. Disentangling contributions of reproductive tract infections to HIV acquisition in African Women. *Sex Transm Dis* 2009;36(6):357-64.
6. Paz-Bailey G, Sternberg M, Puren AJ, et al. Determinants of HIV type 1 shedding from genital ulcers among men in South Africa. *Clin Infect Dis* 2010;50(7):1060-7.
7. Wasserheit JN. Epidemiological synergy. Interrelationships between human immunodeficiency virus infection and other sexually transmitted diseases. *Sex Transm Dis* 1992;19(2):61-77.
8. Mhlongo S, Magooa P, Muller EE, et al. Etiology and STI/HIV coinfections among patients with urethral and vaginal discharge syndromes in South Africa. *Sex Transm Dis* 2010;37(9):566-70.

9. Gray RH, Wawer MJ, Brookmeyer R, et al. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. *Lancet* 2001;357(9263):1149-53.

10. Hester RA, Kennedy SB. Candida infection as a risk factor for HIV transmission. *J Womens Health (Larchmt)* 2003;12(5):487-94.

11. Anorlu R, Imosemi D, Odunukwe N, et al. Prevalence of HIV among women with vaginal discharge in a gynecological clinic. *J Natl Med Assoc* 2004;96(3):367-71.

12. Gray RH, Wawer MJ, Sewankambo NK, et al. Relative risks and population attributable fraction of incident HIV associated with symptoms of sexually transmitted diseases and treatable symptomatic sexually transmitted diseases in Rakai District, Uganda. Rakai Project Team. *AIDS* 1999;13(15):2113-23.

13. McClelland RS, Richardson BA, Hassan WM, et al. Improvement of vaginal health for Kenyan women at risk for acquisition of human immunodeficiency virus type 1: results of a randomized trial. *J Infect Dis* 2008;197(10):1361-8.

14. Frohlich JA, Abdool Karim Q, Mashego MM, et al. Opportunities for treating sexually transmitted infections and reducing HIV risk in rural South Africa. *J Adv Nurs* 2007;60(4):377-83.

15. Kempf MC, Allen S, Zulu I, et al. Enrollment and retention of HIV discordant couples in Lusaka, Zambia. *J Acquir Immune Defic Syndr* 2008;47(1):116-25.

16. World Factbook - Zambia. Central Intelligence Agency; 2009. (Accessed March 3 2011).

17. Johnson LF, Dorrington RE, Bradshaw D, et al. The effect of syndromic management interventions on the prevalence of sexually transmitted infections in South Africa. *Sex Reprod Healthc* 2011;2(1):13-20.

18. Kalichman SC, Cain D, Simbayi LC. Multiple recent sexual partnerships and alcohol use among sexually transmitted infection clinic patients, Cape Town, South Africa. *Sex Transm Dis* 2011;38(1):18-23.

19. Stephenson R, Barker J, Cramer R, et al. The demographic profile of sero-discordant couples enrolled in clinical research in Rwanda and Zambia. *AIDS Care* 2008;20(3):395-405.

20. Modjarrad K, Zulu I, Karita E, et al. Predictors of HIV serostatus among HIV discordant couples in Lusaka, Zambia and female antenatal clinic attendants in Kigali, Rwanda. *AIDS Res Hum Retroviruses* 2005;21(1):5-12.

# TABLES

Table 1. Demographic and behavior characteristics between men and women of heterosexual cohabitating couples by entry into study (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Sex | | | | |
| Men | 1074 (50) | 1113 (50) | 703 (50) | 2890 (50) |
| Women | 1074 (50) | 1113 (50) | 703 (50) | 2890 (50) |
| Age (in years); mean (SD) | | | | |
| Men | 34.09 (8.30) | 35.40 (7.89) | 35.77 (7.56) | 35.00 (8.00) |
| Women | 27.03 (6.78)[2] | 29.03 (6.69)[2] | 29.47 (6.28)[2] | 28.40 (6.71)[2] |
| Age (in years) of first sex, mean (SD) | | | | |
| Men | 17.64 (4.25) | 17.49 (3.59) | 17.68 (3.84) | 17.59 (3.91) |
| Women | 16.27 (2.61)[2] | 16.62 (2.52)[2] | 16.90 (2.53)[2] | 16.56 (2.57)[2] |
| Number of lifetime partners, mean (SD) | | | | |
| Men | 9.59 (9.91) | 13.97 (22.13) | 9.45 (11.42) | 11.15 (7.40) |
| Women | 3.10 (2.55)[2] | 3.66 (7.86)[2] | 3.47 (3.18)[2] | 3.40 (5.27)[2] |
| Number of partners past year, mean (SD) | | | | |
| Men | 1.80 (1.96) | 1.74 (1.82) | 1.52 (1.13) | 1.71 (1.73) |
| Women | 1.12 (0.52)[2] | 1.11 (1.31)[2] | 1.18 (2.38)[2] | 1.13 (1.48)[2] |
| Number of partners past 3 mo, mean (SD) | | | | |
| Men | 1.18 (0.85) | 1.08 (0.62) | 1.07 (0.40) | 1.12 (0.68) |
| Women | 1.01 (0.69)[2] | 0.97 (0.25)[2] | 1.00 (0.35)[2] | 0.99 (0.48)[2] |

Table 1 (cont'd). Demographic and behavior characteristics between men and women of heterosexual cohabitating couples by entry into study (n=5780), Zambia, 1994-2010

| | Study entry 06FEB95-22JAN02 (n=2148) | Study entry 15MAR02-14DEC06 (n=2226) | Study entry 05JAN07-25FEB10 (n=1406) | All Groups (n=5780) |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Sex with other partner[a] | | | | |
| Men | –[1] | 84 (7.55) | 68 (9.67) | 152 (5.26) |
| Women | –[1] | 22 (1.98)[2] | 27 (3.84)[2] | 49 (1.70)[2] |
| Sex with other partner, no condom[a] | | | | |
| Men | 110 (10.24) | –[1] | –[1] | 110 (3.81) |
| Women | 23 (2.14)[2] | –[1] | –[1] | 23 (0.80)[2] |
| Sex with other partner, with condom[a] | | | | |
| Men | 57 (5.31) | –[1] | –[1] | 57 (1.97) |
| Women | 9 (0.84)[2] | –[1] | –[1] | 9 (0.31)[2] |
| Sex with partner, no condom[a] | | | | |
| Men | 857 (79.80) | 551 (49.51) | 563 (80.09) | 1971 (68.20) |
| Women | 842 (78.40) | 888 (79.78)[2] | 584 (83.07) | 2314 (80.07)[3] |
| Sex with partner, with condom[a] | | | | |
| Men | 389 (36.22) | 545 (48.97) | 470 (66.86) | 1404 (48.58) |
| Women | 386 (35.94) | 782 (70.26) | 492 (69.99) | 1660 (57.44) |

[a]Partner refers to main partner in study, past 3 months as reported on Past Medical History

[1]No data/not enough data available

[2]p-value <0.01

[3]p-value <0.05

Table 2. Demographic, behavior and health characteristics by HIV status of heterosexual cohabitating couples by entry into study (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| HIV serostatus | | | | |
| No | 1676 (78.03) | 1860 (83.56) | 1324 (94.17) | 4860 (84.08) |
| Linked | 382 (17.78) | 246 (11.05) | 48 (3.41) | 676 (11.70) |
| Unlinked | 66 (3.07) | 98 (4.40) | 26 (1.85) | 190 (3.28) |
| Unknown linkage | 24 (1.12) | 22 (0.99) | 8 (0.57) | 54 (0.93) |
| Couple HIV status at baseline | | | | |
| Male (+) Female (-) | 1134 (52.79) | 1016 (45.64) | 566 (40.26) | 2716 (46.99) |
| Female (+) Male (-) | 1014 (47.21) | 1210 (54.36) | 840 (59.74) | 3064 (53.01) |
| Age (in years); mean (SD) | | | | |
| HIV - | 30.42 (8.61) | 32.30 (8.22) | 33.20 (8.16) | 31.82 (8.42) |
| HIV + | 30.70 (8.10) | 32.13 (7.73) | 32.05 (7.01)[2] | 31.58 (7.73) |
| Age (in years) of first sex, mean (SD) | | | | |
| HIV - | 16.89 (3.50) | 17.12 (3.08) | 17.51 (3.58) | 17.13 (3.38) |
| HIV + | 17.02 (3.68) | 16.96 (3.17) | 17.07 (2.92)[3] | 17.01 (3.31) |

Table 2 (cont'd). Demographic, behavior and health characteristics by HIV status of heterosexual cohabitating couples by entry into study (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Number of lifetime partners, mean (SD) | | | | |
| HIV - | 5.88 (7.67) | 8.82 (17.81) | 6.38 (9.67) | 7.10 (12.85) |
| HIV + | 6.78 (8.13)[2] | 8.74 (16.86) | 6.56 (8.06) | 7.43 (12.06) |
| Number of partners past year, mean (SD) | | | | |
| HIV - | 1.40 (1.45) | 1.42 (1.55) | 1.38 (2.48) | 1.40 (1.80) |
| HIV + | 1.51 (1.49) | 1.44 (1.68) | 1.32 (0.93) | 1.44 (1.45) |
| Number of partners past 3 mo, mean (SD) | | | | |
| HIV - | 1.06 (0.64) | 1.00 (0.35) | 1.03 (0.31) | 1.03 (0.48) |
| HIV + | 1.12 (0.90) | 1.02 (0.51) | 1.05 (0.44) | 1.07 (0.68)[3] |
| Sex with other partner[a] | | | | |
| HIV - | –[1] | 49 (4.40) | 39 (5.55) | 88 (3.04) |
| HIV + | –[1] | 57 (5.12) | 56 (7.97) | 113 (3.91) |
| Sex with other partner, no condom[a] | | | | |
| HIV - | 59 (5.49) | –[1] | –[1] | 59 (2.04) |
| HIV + | 74 (6.89) | –[1] | –[1] | 74 (2.56) |
| Sex with other partner, with condom[a] | | | | |
| HIV - | 27 (2.51) | –[1] | –[1] | 27 (0.93) |
| HIV + | 39 (3.63) | –[1] | –[1] | 39 (1.35) |

Table 2 (cont'd). Demographic, behavior and health characteristics by HIV status of heterosexual cohabitating couples by entry into study (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Sex with partner, no condom[a] | | | | |
| HIV - | 848 (78.96) | 691 (62.08) | 564 (80.23) | 2103 (72.77) |
| HIV + | 851 (79.24) | 748 (67.21) | 583 (82.93) | 2182 (75.50) |
| Sex with partner, with condom[a] | | | | |
| HIV - | 386 (35.94) | 642 (57.68) | 478 (67.99) | 1506 (52.11) |
| HIV + | 389 (36.22) | 685 (61.55) | 484 (68.85) | 1558 (53.91) |

[a]Partner refers to main partner in study, past 3 months as reported on Past Medical History

[1]No data/not enough data available

[2]p-value <0.01

[3]p-value <0.05

Table 3. Baseline characteristics between men and women of heterosexual cohabitating couples by medical form and entry into study (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Past Medical History | | | | |
| Gonorrhea[a] | | | | |
| Men | 201 (18.72) | 21 (1.89) | 14 (1.99) | 236 (8.17) |
| Women | 29 (2.70)[2] | 8 (0.72)[3] | 2 (0.28)[2] | 39 (1.35)[2] |
| Syphilis[a] | | | | |
| Men | 124 (11.55) | 182 (16.35) | 44 (6.26) | 350 (12.11) |
| Women | 163 (15.18)[2] | 274 (24.62)[2] | 46 (6.54) | 483 (16.71)[2] |
| Ulceration[ab] | | | | |
| Men | 206 (19.18) | 279 (25.07) | 74 (10.53) | 559 (19.34) |
| Women | 88 (8.19)[2] | 132 (11.86)[2] | 15 (2.13)[2] | 235 (8.13)[2] |
| Discharge[ac] | | | | |
| Men | 119 (11.08) | 51 (4.58) | 26 (3.70) | 196 (6.78) |
| Women | 175 (16.29)[2] | 106 (9.52)[2] | 4 (0.57)[2] | 285 (9.86)[2] |
| Cystitis/dysuria[a] | | | | |
| Men | 207 (19.27) | 108 (9.70) | 40 (5.69) | 355 (12.28) |
| Women | 139 (12.94)[2] | 78 (7.01)[3] | 14 (1.99)[2] | 231 (7.99)[2] |
| Lower abdominal pain[a] | | | | |
| Men | 179 (16.67) | 121 (10.87) | –[1] | 300 (10.38) |
| Women | 209 (19.46)[2] | 107 (9.61) | 7 (1.00) | 323 (11.18)[2] |

Table 3 (cont'd). Baseline characteristics between men and women of heterosexual cohabitating couples by medical form and entry into study (n=5780), Zambia, 1994-2010

| | Study entry 06FEB95-22JAN02 (n=2148) | Study entry 15MAR02- 14DEC06 (n=2226) | Study entry 05JAN07-25FEB10 (n=1406) | All Groups (n=5780) |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Physical Examination | | | | |
| Inflammation[d] | | | | |
| Men | 27 (2.51) | 5 (0.45) | 6 (0.85) | 38 (1.31) |
| Women | 80 (7.45)[2] | 14 (1.26)[3] | 2 (0.28) | 96 (3.32)[2] |
| Ulceration[e] | | | | |
| Men | 72 (6.70) | 33 (2.96) | 18 (2.56) | 123 (4.26) |
| Women | 57 (5.31) | 35 (3.14) | 4 (0.57)[2] | 96 (3.32) |
| Discharge[f] | | | | |
| Men | 36 (3.35) | 9 (0.81) | 8 (1.14) | 53 (1.83) |
| Women | 154 (14.34)[2] | 137 (12.31)[2] | 3 (0.43) | 294 (10.17)[2] |
| Inflammation composite[g] | | | | |
| Men | 57 (5.31) | –[1] | –[1] | 57 (1.97) |
| Women | 208 (19.37)[2] | 146 (13.12) | 5 (0.71) | 359 (12.42)[2] |
| Ulceration composite[h] | | | | |
| Men | 72 (6.70) | 33 (2.96) | 18 (2.56) | 123 (4.26) |
| Women | 124 (11.55)[2] | 40 (3.59) | 6 (0.85)[3] | 170 (5.88)[2] |
| Laboratory | | | | |
| Gonorrhea[i] | | | | |
| Men | 1 (0.09) | –[1] | –[1] | 1 (0.03) |
| Women | 28 (2.61)[2] | –[1] | –[1] | 28 (0.97)[2] |

Table 3 (cont'd). Baseline characteristics between men and women of heterosexual cohabitating couples by medical form and entry into study (n=5780), Zambia, 1994-2010

| | Study entry 06FEB95-22JAN02 (n=2148) | Study entry 15MAR02-14DEC06 (n=2226) | Study entry 05JAN07-25FEB10 (n=1406) | All Groups (n=5780) |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Syphilis[j] | | | | |
| Men | 187 (17.41) | 187 (16.80) | 50 (7.11) | 424 (14.67) |
| Women | 264 (24.58)[2] | 249 (22.37)[2] | 53 (7.54) | 566 (19.58)[2] |
| Herpes Simplex Virus-2[k] | | | | |
| Men | –[1] | 36 (3.23) | 2 (0.28) | 38 (1.31) |
| Women | –[1] | 74 (6.65)[2] | –[1] | 74 (2.56)[2] |
| Candida[l] | | | | |
| Men | 1 (0.09) | –[1] | –[1] | 1 (0.03) |
| Women | 142 (13.22)[2] | –[1] | –[1] | 142 (4.91)[2] |
| Bacterial vaginosis[m] | | | | |
| Men | –[1] | –[1] | –[1] | –[1] |
| Women | 201 (18.72) | –[1] | –[1] | 201 (6.96) |
| Trichomonas[n] | | | | |
| Men | 1 (0.09) | –[1] | –[1] | 1 (0.03) |
| Women | 208 (19.37)[2] | 134 (12.04) | 36 (5.12) | 378 (13.08)[2] |

[a]Patient report for past 3 months or year

[b]Acute or chronic/recurrent genital ulcer

[c]Vaginal or urethral discharge

[d]Inflammation, cervical inflammaion, vaginal inflammation

[e]Genital ulceration, cervical ulceration, vaginal ulceration

[f]Urethral, cervical and vaginal discharge

[g]Inflammation composite variable including inflammation, cervical and vaginal inflammation and cervical, vaginal and urethral discharge

[h]Ulceration composite variable including external ulceration, cervical and genital ulceration and cervical and vaginal erosion/friability

[i]Gonorrhea lab test by gram stain or culture

[j]Syphilis lab test by Rapid Plasma Reagin (RPR)

[k]Herpes Simplex Virus-2 lab test by HerpeSelect

[l]Candida lab test by wet prep or gram cervix

[m]Bacterial vaginosis lab test by wet prep

[n]Trichomonas lab test by wet prep

–[l]No data/not enough data

[2]p-value <0.01

[3]p-value <0.05

Table 4. Baseline characteristics by HIV status of heterosexual cohabitating couples by medical form and entry into study (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Past Medical History | | | | |
| Gonorrhea[a] | | | | |
| HIV - | 96 (8.94) | 12 (1.08) | 8 (1.14) | 116 (4.01) |
| HIV + | 134 (12.48)[2] | 17 (1.53) | 8 (1.14) | 159 (5.50)[2] |
| Syphilis[a] | | | | |
| HIV - | 107 (9.96) | 176 (15.81) | 40 (5.69) | 323 (11.18) |
| HIV + | 180 (16.76)[2] | 280 (25.16)[2] | 50 (7.11) | 510 (17.65)[2] |
| Ulceration[ab] | | | | |
| HIV - | 98 (9.12) | 126 (11.32) | 26 (3.70) | 250 (8.65) |
| HIV + | 196 (18.25)[2] | 285 (25.61)[2] | 63 (8.96)[2] | 544 (18.82)[2] |
| Discharge[ac] | | | | |
| HIV - | 135 (12.57) | 62 (5.57) | 17 (2.42) | 214 (7.40) |
| HIV + | 159 (14.80) | 95 (8.54)[2] | 13 (1.85) | 267 (9.24)[3] |
| Cystitis/dysuria[a] | | | | |
| HIV - | 132 (12.29) | 78 (7.01) | 24 (3.41) | 234 (8.10) |
| HIV + | 214 (19.93)[2] | 108 (9.70)[3] | 30 (4.27) | 352 (12.18)[2] |
| Lower abdominal pain[a] | | | | |
| HIV - | 159 (14.80) | 102 (9.16) | 2 (0.28) | 263 (9.10) |
| HIV + | 229 (21.32)[2] | 126 (11.32) | 5 (0.71) | 360 (12.46)[2] |

Table 4 (cont'd). Baseline characteristics by HIV status of heterosexual cohabitating couples by medical form and entry into study (n=5780), Zambia, 1994-2010

| | Study entry 06FEB95-22JAN02 (n=2148) | Study entry 15MAR02-14DEC06 (n=2226) | Study entry 05JAN07-25FEB10 (n=1406) | All Groups (n=5780) |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Physical Examination | | | | |
| Inflammation[d] | | | | |
| HIV - | 44 (4.10) | 3 (0.27) | 3 (0.43) | 50 (1.73) |
| HIV + | 63 (5.87) | 16 (1.44)[2] | 5 (0.71) | 84 (2.91)[2] |
| Ulceration[e] | | | | |
| HIV - | 34 (3.17) | 19 (1.71) | 3 (0.43) | 56 (1.94) |
| HIV + | 95 (8.85)[2] | 49 (4.40)[2] | 19 (2.70)[2] | 163 (5.64)[2] |
| Discharge[f] | | | | |
| HIV - | 92 (8.57) | 58 (5.21) | 5 (0.71) | 155 (5.36) |
| HIV + | 98 (9.12) | 88 (7.91)[3] | 6 (0.85) | 192 (6.64)[3] |
| Inflammation composite[g] | | | | |
| HIV - | 127 (11.82) | 57 (5.12) | 2 (0.28) | 186 (6.44) |
| HIV + | 138 (12.85) | 89 (8.00) | 3 (0.43) | 230 (7.96) |
| Ulceration composite[h] | | | | |
| HIV - | 66 (6.15) | 21 (1.89) | 5 (0.71) | 92 (3.18) |
| HIV + | 130 (12.10)[2] | 52 (4.67)[2] | 19 (2.70)[2] | 201 (6.96)[2] |
| Laboratory | | | | |
| Gonorrhea[i] | | | | |
| HIV - | 10 (0.93) | –[1] | –[1] | 10 (0.35) |
| HIV + | 19 (1.77) | –[1] | –[1] | 19 (0.66) |

Table 4 (cont'd). Baseline characteristics by HIV status of heterosexual cohabitating couples by medical form and entry into study (n=5780), Zambia, 1994-2010

| | Study entry 06FEB95-22JAN02 (n=2148) | Study entry 15MAR02-14DEC06 (n=2226) | Study entry 05JAN07-25FEB10 (n=1406) | All Groups (n=5780) |
|---|---|---|---|---|
| **Characteristic** | **n (%)** | **n (%)** | **n (%)** | **n (%)** |
| Syphilis[j] | | | | |
| HIV - | 180 (16.76) | 156 (14.02) | 46 (6.54) | 382 (13.22) |
| HIV + | 271 (25.23)[2] | 280 (25.16)[2] | 57 (8.11) | 608 (21.04)[2] |
| Herpes Simplex Virus-2[k] | | | | |
| HIV - | –[1] | 1 (0.09) | 1 (0.14) | 2 (0.07) |
| HIV + | –[1] | 109 (9.79)[2] | 1 (0.14) | 110 (3.81)[3] |
| Candida[l] | | | | |
| HIV - | 71 (6.61) | –[1] | –[1] | 71 (2.46) |
| HIV + | 72 (6.70) | –[1] | –[1] | 72 (2.49) |
| Bacterial vaginosis[m] | | | | |
| HIV - | 98 (9.12) | –[1] | –[1] | 98 (3.39) |
| HIV + | 103 (9.59) | –[1] | –[1] | 103 (3.56) |
| Trichomonas[n] | | | | |
| HIV - | 92 (8.57) | 56 (5.03) | 11 (1.56) | 159 (5.50) |
| HIV + | 117 (10.89)[2] | 78 (7.01) | 25 (3.56) | 220 (7.61)[3] |

[a]Patient report for past 3 months or year

[b]Acute or chronic/recurrent genital ulcer

[c]Vaginal or urethral discharge

[d]Inflammation, cervical inflammaion, vaginal inflammation

[e]Genital ulceration, cervical ulceration, vaginal ulceration

[f]Urethral, cervical and vaginal discharge

[g]Inflammation composite variable including inflammation, cervical and vaginal inflammation and cervical, vaginal and urethral discharge

[h]Ulceration composite variable including external ulceration, cervical and genital ulceration and cervical and vaginal erosion/friability

[i]Gonorrhea lab test by gram stain or culture

[j]Syphilis lab test by Rapid Plasma Reagin (RPR)

[k]Herpes Simplex Virus-2 lab test by HerpeSelect

[l]Candida lab test by wet prep or gram cervix

[m]Bacterial vaginosis lab test by wet prep

[n]Trichomonas lab test by wet prep

–[1]No data/not enough data

[2]p-value <0.01

[3]p-value <0.05

Table 5. Univariate odds ratios of clinical and behavioral characteristics with HIV prevalence at baseline (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **cOR (95% CI)** | **cOR (95% CI)** | **cOR (95% CI)** | **cOR (95% CI)** |
| Past Medical History | | | | |
| Gonorrhea[a] | 1.35 (1.01, 1.79)[3] | 1.55 (0.73, 3.26) | 1.36 (0.50, 3.67) | 1.54 (1.20, 1.98)[2] |
| Syphilis[a] | 1.84 (1.43, 2.38)[2] | 1.73 (1.40, 2.13)[2] | 1.27 (0.82, 1.97) | 1.66 (1.43, 1.94)[2] |
| Ulceration[ab] | 2.14 (1.64, 2.78)[2] | 3.06 (2.42, 3.87)[2] | 3.63 (2.24, 5.88)[2] | 2.67 (2.27, 3.15)[2] |
| Discharge[ac] | 1.24 (0.97, 1.59) | 1.49 (1.07, 2.10)[3] | 1.02 (0.49, 2.14) | 1.24 (1.03, 1.50)[3] |
| Cystitis/dysuria[a] | 1.74 (1.37, 2.20)[2] | 1.48 (1.09, 2.00)[3] | 1.56 (0.89, 2.73) | 1.63 (1.37, 1.94)[2] |
| Lower abdominal pain[a] | 1.58 (1.26, 1.97)[2] | 1.28 (0.97, 1.69) | 1.69 (0.33, 8.77) | 1.36 (1.15, 1.61)[2] |
| Physical Examination | | | | |
| Inflammation[d] | 1.55 (1.04, 2.31)[3] | 5.01 (1.45, 17.28)[3] | 2.07 (0.49, 8.87) | 1.61 (1.13, 2.30)[2] |
| Ulceration[e] | 2.93 (1.96, 4.38)[2] | 2.66 (1.55, 4.55)[2] | 8.76 (2.56, 29.95)[2] | 3.08 (2.27, 4.20)[2] |
| Discharge[f] | 1.15 (0.85, 1.57) | 1.35 (0.95, 1.92) | 1.47 (0.44, 4.91) | 1.16 (0.93, 1.45) |
| Inflammation composite[g] | 1.19 (0.91, 1.55) | 1.36 (0.96, 1.95) | 1.02 (0.17, 6.12) | 1.11 (0.90, 1.36) |
| Ulceration composite[h] | 2.19 (1.60, 2.99)[2] | 2.52 (1.51, 4.22)[2] | 4.95 (1.82, 13.47)[2] | 2.24 (1.74, 2.88)[2] |
| Laboratory | | | | |
| Gonorrhea[i] | 2.14 (0.98, 4.65) | –[1] | –[1] | 2.15 (0.99, 4.67) |
| Syphilis[j] | 1.72 (1.39, 2.13)[2] | 2.04 (1.64, 2.54)[2] | 1.25 (0.83, 1.89) | 1.72 (1.50, 1.98)[2] |
| Herpes Simplex Virus-2[k] | –[1] | 21.85 (1.69, 282.42)[3] | –[1] | 6.71 (0.82, 54.91) |
| Candida[l] | 1.12 (0.78, 1.59) | –[1] | –[1] | 1.12 (0.79, 1.60) |
| Bacterial vaginosis[m] | 1.18 (0.86, 1.61) | –[1] | –[1] | 1.18 (0.87, 1.62) |
| Trichomonas[n] | 1.51 (1.11, 2.06)[2] | 1.19 (0.83, 1.72) | 1.60 (0.77, 3.32) | 1.26 (1.01, 1.57)[3] |

Table 5 (cont'd). Univariate odds ratios of clinical and behavioral characteristics with HIV prevalence at baseline (n=5780), Zambia, 1994-2010

| | Study entry 06FEB95-22JAN02 (n=2148) | Study entry 15MAR02-14DEC06 (n=2226) | Study entry 05JAN07-25FEB10 (n=1406) | All Groups (n=5780) |
|---|---|---|---|---|
| **Characteristic** | **cOR (95% CI)** | **cOR (95% CI)** | **cOR (95% CI)** | **cOR (95% CI)** |
| Sexual behavior | | | | |
| Sex with other partner | –[1] | 1.24 (0.83, 1.85) | 1.82 (1.18, 2.90)[2] | 1.50 (1.12, 2.02)[2] |
| Sex with other partner, no condom | 1.21 (0.84, 1.73) | –[1] | –[1] | 1.21 (0.84, 1.73) |
| Sex with other partner, with condom | 1.38 (0.84, 2.28) | –[1] | –[1] | 1.38 (0.84, 2.28) |
| Sex with partner, no condom | 0.99 (0.80, 1.23) | 1.05 (0.85, 1.31) | 1.17 (0.89, 1.55) | 1.06 (0.93, 1.22) |
| Sex with partner, with condom | 1.01 (0.85, 1.20) | 1.02 (0.83, 1.24) | 1.02 (0.81, 1.29) | 1.03 (0.92, 1.14) |

[a]Patient report for past 3 months or year

[b]Acute or chronic/recurrent genital ulcer

[c]Vaginal or urethral discharge

[d]Inflammation, cervical inflammaion, vaginal inflammation

[e]Genital ulceration, cervical ulceration, vaginal ulceration

[f]Urethral, cervical and vaginal discharge

[g]Inflammation composite variable including inflammation, cervical and vaginal inflammation and cervical, vaginal and urethral discharge

[h]Ulceration composite variable including external ulceration, cervical and genital ulceration and cervical and vaginal erosion/friability

[i]Gonorrhea lab test by gram stain or culture

[j]Syphilis lab test by Rapid Plasma Reagin (RPR)

[k]Herpes Simplex Virus-2 lab test by HerpeSelect

[l]Candida lab test by wet prep or gram cervix

[m]Bacterial vaginosis lab test by wet prep

[n]Trichomonas lab test by wet prep

[1]Not enough information for OR calcuation
[2]p-value <0.01
[3]p-value <0.05

Table 6. Multivariable odds ratios of clinical and behavioral characteristics with HIV prevalence at baseline (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **aOR (95% CI)** | **aOR (95% CI)** | **aOR (95% CI)** | **aOR (95% CI)** |
| Past Medical History | | | | |
| Gonorrhea[a] | 2.65 (1.06, 6.65)[3] | | | 2.46 (1.07, 5.66)[3] |
| Syphilis[a] | | | | |
| Ulceration[ab] | | | 3.06 (1.84, 5.08)[2] | 1.98 (1.42, 2.76)[2] |
| Discharge[ac] | | | | |
| Cystitis/dysuria[a] | 1.56 (1.02, 2.38)[3] | | | 1.40 (1.01, 1.94)[3] |
| Lower abdominal pain[a] | 1.56 (1.09, 2.24)[2] | | | |
| Physical Examination | | | | |
| Inflammation[d] | | | | |
| Ulceration[e] | | | 4.79 (1.35, 17.00)[3] | |
| Discharge[f] | | | | |
| Inflammation composite[g] | | | | |
| Ulceration composite[h] | | | | |
| Laboratory | | | | |
| Gonorrhea[i] | | | | |
| Syphilis[j] | | | | 1.77 (1.38, 2.27)[2] |
| Herpes Simplex Virus-2[k] | | 13.21 (1.08, 161.98)[3] | | |
| Candida[l] | | | | |
| Bacterial vaginosis[m] | | | | |
| Trichomonas[n] | | | | |

Table 6 (cont'd). Multivariable odds ratios of clinical and behavioral characteristics with HIV prevalence at baseline (n=5780), Zambia, 1994-2010

| | **Study entry 06FEB95-22JAN02 (n=2148)** | **Study entry 15MAR02-14DEC06 (n=2226)** | **Study entry 05JAN07-25FEB10 (n=1406)** | **All Groups (n=5780)** |
|---|---|---|---|---|
| **Characteristic** | **aOR (95% CI)** | **aOR (95% CI)** | **aOR (95% CI)** | **aOR (95% CI)** |
| Sexual behavior | | | | |
| Sex with other partner | | | | |
| Sex with other partner, no condom | | | | |
| Sex with other partner, with condom | | | | |
| Sex with partner, no condom | | | | |
| Sex with partner, with condom | | | | |

[a]Patient report for past 3 months or year

[b]Acute or chronic/recurrent genital ulcer

[c]Vaginal or urethral discharge

[d]Inflammation, cervical inflammaion, vaginal inflammation

[e]Genital ulceration, cervical ulceration, vaginal ulceration

[f]Urethral, cervical and vaginal discharge

[g]Inflammation composite variable including inflammation, cervical and vaginal inflammation and cervical, vaginal and urethral discharge

[h]Ulceration composite variable including external ulceration, cervical and genital ulceration and cervical and vaginal erosion/friability

[i]Gonorrhea lab test by gram stain or culture

[j]Syphilis lab test by Rapid Plasma Reagin (RPR)

[k]Herpes Simplex Virus-2 lab test by HerpeSelect

[l]Candida lab test by wet prep or gram cervix

[m]Bacterial vaginosis lab test by wet prep

[n]Trichomonas lab test by wet prep

[2]p-value <0.01

[3]p-value <0.05

# APPENDIX A

## Annotated SAS Code

```
************************************;
*****   ALEXANDRA VAIA RICCA    *****;
*****  MASTER OF PUBLIC HEALTH  *****;
*****    GLOBAL EPIDEMIOLOGY    *****;
*****          THESIS           *****;
*****           2011            *****;
************************************;

libname ricca 'H:\Thesis';

data thesis;
set ricca.stibase31mar11;
run;

*Exploring data;
proc freq data=thesis;
tables n_tpha;
run;

proc freq data=thesis;
tables hsvrsult hsvrsult2;
run;

proc freq data=thesis;
tables n_candwp n_cancx n_bactwp;
run;

proc freq data=thesis;
tables hvagin hureth;
run;

proc freq data=thesis;
tables hcysd;
run;

proc freq data=thesis;
tables hapain;
run;

proc freq data=thesis;
tables pinfge paingi pbingi;
run;

proc freq data=thesis;
tables pexgul paulce pbulce;
run;

proc freq data=thesis;
tables pdisch padisc pbdisc;
run;
```

```
************************************;
***** Cleaning and new variables *****;
************************************;

data cleanthesis;
set work.thesis;

*Create variable CJ treating "D" as "-";
if cjbase='D+' then cj='-+';
if cjbase='-+' then cj='-+';
if cjbase='+-' then cj='+-';
if cjbase='+D' then cj='+-';

*Delete couples where womanage not between 16 and 45 and/or manage not
between 16 and 65;
if htid=265 then delete;
if htid=120 then delete;
if htid=130 then delete;
if htid=63 then delete;
if htid=66 then delete;
if htid=16 then delete;
if htid=992 then delete;
if htid=323 then delete;
if htid=352 then delete;
if htid=280 then delete;
if htid=286 then delete;
if htid=3083 then delete;
if htid=3272 then delete;
if htid=781 then delete;
if htid=460 then delete;
if htid=32 then delete;
if htid=101 then delete;

*Create variable for age;
if sex='F' and womanage>0 then age=womanage;
else if sex='M' and manage>0 then age=manage;
else age=.;
label age='Age of client';

*Create variable for sex since it is character;
if sex='M' then sexnum=0;
else if sex='F' then sexnum=1;
label sexnum='Client Sex';

*Create variable that gives baseline HIV prevalence;
*HIV prevalence variable for M;
if cj='-+' and sex='F' then HIVbase=1;
else if cj='+-' and sex='M' then HIVbase=1;
else HIVbase=0;
label HIVbase='HIV baseline status';

*Create variable for seroconversion yes/no for HIV incidence;
*Seroconversion variable for M;
if serostatMAR10=1 and cj='-+' and sex='M' then seroconvert=1;
else if serostatMAR10=2 and cj='-+' and sex='M' then seroconvert=1;
else if serostatMAR10=15 and cj='-+' and sex='M' then seroconvert=1;
*Seroconversion variable for F;
```

```
else if serostatMAR10=1 and cj='+-' and sex='F' then seroconvert=1;
else if serostatMAR10=2 and cj='+-' and sex='F' then seroconvert=1;
else if serostatMAR10=15 and cj='+-' and sex='F' then seroconvert=1;
else seroconvert=0;
label seroconvert='HIV seroconversion';

*Create variable htidgroup for 3 time period groups;
if htid le 1634 then htidgroup=1;
else if htid le 2994 then htidgroup=2;
else htidgroup=3;
label htidgroup='HTID group';

*SEXUAL BEHAVIOR VARIABLES;
*Create variable for sex outside project partner yes/no;
if hsxot ge 1 then othersex=1;
else if hsxot=0 then othersex=0;
else othersex=.;
label othersex='Sex with someone other than project partner';

*Create variable for sex outside project partner without condom yes/no;
if f_sxotnc ge 1 then othersexnc=1;
else if f_sxotnc=0 then othersexnc=0;
else othersexnc=.;
label othersexnc='Sex with someone other than project partner without
condom';

*Create variable for sex outside project partner with condom yes/no;
if f_sxotwc ge 1 then othersexwc=1;
else if f_sxotwc=0 then othersexwc=0;
else othersexwc=.;
label othersexwc='Sex with someone other than project partner with
condom';

*Create variable for sex with partner in project without condom yes/no;
if hsxprnc ge 1 then nocondomsex=1;
else if hsxprnc=0 then nocondomsex=0;
else nocondomsex=.;
label nocondomsex='Sex with partner in project without condom in past 3
months';

*Create variable for sex with partner in project using condom yes/no;
if hsxprwc ge 1 then condomsex=1;
else if hsxprwc=0 then condomsex=0;
else condomsex=.;
label condomsex='Sex with partner in project using condom in past 3
months';

*Create age at first sex variable, not including "0" or "99";
if hagesex=99 or hagesex lt 8 then firstsex=.;
else if hagesex ge 8 then firstsex=hagesex;
else firstsex=.;
label firstsex='Age at first sex';

*Create number of lifetime partners variable, not including 999;
if hlifsex=999 then lifesex=.;
else if hlifsex ge 0 then lifesex=hlifsex;
else lifesex=.;
```

```
label lifesex='Number of lifetime partners';

*Create number of partners in past year variable, not including 999;
if hlyrsex=999 then yearsex=.;
else if hlyrsex ge 0 then yearsex=hlyrsex;
else yearsex=.;
label yearsex='Number of partners in the past year';

*Create number of partners in past 3 months variable, not including
901;
if h3mosex=901 then sex3mo=.;
else if h3mosex ge 0 then sex3mo=h3mosex;
else sex3mo=.;
label sex3mo='Number of partners in the past 3 months';

*LABORATORY VARIABLES;
*Recode N_RPR and N_TPHA variable to:
rpr=0=Negative
rpr=1=Positive;
if n_rpr=1 or n_rpr=11 or n_rpr=12 or n_rpr=14 or n_rpr=18 or n_rpr=116
   or n_rpr=128 or n_rpr=164 or n_rpr=1128 or n_rpr=1256 or n_rpr=1512
then rpr=1;
else if n_rpr=2 then rpr=0;
else rpr=.;
label rpr='Syphilis lab test';

*Recode N_TPHA variable to:
tpha=0=Negative
tpha=1=positive;
if n_tpha ge 128 or n_tpha='P' then tpha=1;
else if n_tpha='N' then tpha=0;
else tpha=.;
label tpha='Syphilis confirm lab test';

*Combine syphilis lab tests to:
syphlab=0=negative
syphlab=1=positive;
if n_rpr=1 or n_rpr=11 or n_rpr=12 or n_rpr=14 or n_rpr=18 or n_rpr=116
or
   n_rpr=128 or n_rpr=164 or n_rpr=1128 or n_rpr=1256 or n_rpr=1512 or
   n_tpha ge 128 or n_tpha='P' then syphlab=1;
else if n_rpr=2 or n_tpha='N' then syphlab=0;
else syphlab=.;
label syphlab='Syphilis combined lab test';

*Recode N_GRAM and N_CULTUR variable to:
gontest=0=negative
gontest=1=positive;
if n_gram='+' or n_cultur='+' then gontest=1;
else if n_gram='-' or n_cultur='-' then gontest=0;
else gontest=.;
label gontest='Gonorrhea lab test';

*Recode HSVrsult and HSVrsult2 to:
HSV=0=negative, doubtful
HSV=1=positive;
if HSVrsult='P' or HSVrsult2='P' then hsv=1;
```

```
else if HSVrsult='N' or HSVrsult='D' or HSVrsult2='N' or HSVrsult2='D'
then hsv=0;
else hsv=.;
label hsv='HSV-2 lab test';

*Recode n_candwp and n_cancx to:
candida=0=negative
candida=1=positive;
if n_candwp='+' or n_cancx='+' then candida=1;
else if n_candwp='-' or n_cancx='-' then candida=0;
else candida=.;
label candida='Candida lab test';

*Recode n_bactwp to:
bacvag=0=negative
bacvag=1=positive;
if n_bactwp='+' then bacvag=1;
else if n_bactwp='-' then bacvag=0;
else bacvag=.;
label bacvag='Bacterial vaginosis lab test';

*Recode N_TRICWP variable to:
trichtest=0=negative
trichtest=1=positive;
if n_tricwp='+' then trichtest=1;
else if n_tricwp='-' then trichtest=0;
else trichtest=.;
label trichtest='Trich lab test';

*PAST MEDICAL HISTORY VARIABLES;
*Create gonoreport variable for patient report of gonorrhea (H09GONO,
H47GONO, HGONO) in the past:
gonoreport=0=negative
gonoreport=1=positive;
if h09gono ge 1 or h47gono ge 1 or hgono=1 then gonoreport=1;
else if h09gono=0 or h47gono=0 or hgono ge 2 then gonoreport=0;
else gonoreport=.;
label gonoreport='Patient reports gonorrhea in the past';

*Create syphreport variable for patient report of syphilis (H10SYPH,
H48SYPH, HSYPH) in the past:
syphreport=0=negative
syphreport=1=positive;
if h10syph ge 1 or h48syph ge 1 or hsyph=1 then syphreport=1;
else if h10syph=0 or h48syph=0 or hsyph ge 2 then syphreport=0;
else syphreport=.;
label syphreport='Patient reports syphilis in the past';

*Create ulcerreport variable for patient report of ulcer (H11ULCER,
H49ULCER, HACGEUL, HCHGEUL) in the past:
ulcerreport=0=negative
ulcereport=1=positive;
if h11ulcer ge 1 or h49ulcer ge 1 or hacgeul=1 or hchgeul=1 then
ulcerreport=1;
else if h11ulcer=0 or h49ulcer=0 or hacgeul ge 2 or hchgeul ge 2 then
ulcerreport=0;
else ulcerreport=.;
```

```
label ulcerreport='Patient reports ulcer in the past';

*Create discreport variable for patient report of discharge (Hvagin,
Hureth) in the past:
discreport=0=negative
discreport=1=positive;
if hvagin=1 or hureth=1 then discreport=1;
else if hvagin=0 or hvagin=2 or hvagin=3 or
        hureth=0 or hureth=2 or hureth=3 then discreport=0;
else discreport=.;
label discreport='Patient reports discharge in the past';

*Create cysd variable for patient report of cystitis/dysuria (Hcysd);
if hcysd=1 then cysdreport=1;
else if hcysd=0 or hcysd=2 or hcysd=3 then cysdreport=0;
else cysdreport=.;
label cysdreport='Patient reports cystitis/dysuria in the past';

*Create lapainreport variable for patient report of lower abdominal
pain (hlapain);
if hapain=1 then lapainreport=1;
else if hapain=0 or hapain=2 or hapain=3 then lapainreport=0;
else lapainreport=.;
label lapainreport='Patient reports lower ab pain in the past';

*PHYSICAL EXAM VARIABLES;
*Create infcomp composite inflammation variable (separate from pgenif)
with pinfge, paingi, pbingi;
if pinfge=1 or paingi=1 or pbingi=1 then infcomp=1;
else if pinfge=0 or pinfge=2 or
        paingi=0 or paingi=2 or paingi=9 or
        pbingi=0 or pbingi=2 or pbingi=9 then infcomp=0;
else infcomp=.;
label infcomp='Inflammation';

*Create ulcercomp composite ulceration variable (separate from pgnulc)
with pexgul, paulce, pbulce;
if pexgul=1 or paulce=1 or pbulce=1 then ulcercomp=1;
else if pexgul=0 or pexgul=2 or
        paulce=0 or paulce=2 or paulce=9 or
        pbulce=0 or pbulce=2 or pbulce=9 then ulcercomp=0;
else ulcercomp=.;
label ulcercomp='Ulceration';

*Create discomp composite discharge variable with pdisch, padisc,
pbdisc;
if pdisch=1 or padisc=1 or pbdisc=1 then discomp=1;
else if pdisch=0 or pdisch=2 or pdisch=9 or
        padisc=0 or padisc=2 or padisc=9 or
            pbdisc=0 or pbdisc=2 or pbdisc=9 then discomp=0;
else discomp=.;
label discomp='Discharge';

run;
```

```
*------------------------;
* Information for tables ;
*------------------------;

proc freq data=cleanthesis;
tables cj/missing;
run;

proc freq data=cleanthesis;
tables serostatMAR10/missing;
run;

proc freq data=cleanthesis;
tables hivbase;
run;

proc freq data=cleanthesis;
tables seroconvert;
where hivbase=0;
run;

proc freq data=cleanthesis;
tables seroconvert;
where sex='F' and hivbase=0;
run;

proc freq data=cleanthesis;
tables seroconvert;
where sex='M' and hivbase=0;
run;

proc ttest data=cleanthesis;
class sex;
var seroconvert;
where hivbase=0;
run;

proc sort data=cleanthesis;
by htidgroup;
run;

proc freq data=cleanthesis;
tables sex;
by htidgroup;
run;

proc freq data=cleanthesis;
tables cj;
where htidgroup=1;
run;

proc freq data=cleanthesis;
tables cj;
where htidgroup=2;
run;

proc freq data=cleanthesis;
```

```
tables cj;
where htidgroup=3;
run;

proc freq data=cleanthesis;
tables serostatMAR10;
by htidgroup;
run;

proc freq data=cleanthesis;
tables cjbase;
run;

proc freq data=cleanthesis;
tables sex;
run;

proc freq data=cleanthesis;
tables f_sxotnc f_sxotwc;
run;

proc freq data=cleanthesis;
tables age;
where sex='F';
run; *Female age range=16-45;

proc freq data=cleanthesis;
tables age;
where sex='M';
run; *Male age range=19-65;

proc ttest data=cleanthesis;
class sex;
var age;
run;

proc univariate data=cleanthesis;
class sex;
var age;
run;

proc univariate data=cleanthesis;
class sex;
var age;
where htidgroup=1;
run;

proc univariate data=cleanthesis;
class sex;
var age;
where htidgroup=2;
run;

proc univariate data=cleanthesis;
class sex;
var age;
where htidgroup=3;
```

```
run;

proc ttest data=cleanthesis;
class sex;
var age;
run;

proc ttest data=cleanthesis;
class sex;
var age;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class sex;
var age;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class sex;
var age;
where htidgroup=3;
run;

*By HIV status;
proc univariate data=cleanthesis;
class hivbase;
var age;
run;

proc univariate data=cleanthesis;
class hivbase;
var age;
where htidgroup=1;
run;

proc univariate data=cleanthesis;
class hivbase;
var age;
where htidgroup=2;
run;

proc univariate data=cleanthesis;
class hivbase;
var age;
where htidgroup=3;
run;

proc ttest data=cleanthesis;
class hivbase;
var age;
run;

proc ttest data=cleanthesis;
class hivbase;
var age;
```

```
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var age;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class hivbase;
var age;
where htidgroup=3;
run;

*Age at first sex;
*By sex;
proc ttest data=cleanthesis;
class sex;
var firstsex;
run;

proc ttest data=cleanthesis;
class sex;
var firstsex;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class sex;
var firstsex;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class sex;
var firstsex;
where htidgroup=3;
run;

proc univariate data=cleanthesis;
class sex;
var firstsex;
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=1 and sex='M';
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=1 and sex='F';
run;

proc univariate data=cleanthesis;
var firstsex;
```

```
where htidgroup=2 and sex='M';
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=2 and sex='F';
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=3 and sex='M';
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=3 and sex='F';
run;

*By HIV status;
proc univariate data=cleanthesis;
class hivbase;
var firstsex;
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=1 and hivbase=0;
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=1 and hivbase=1;
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=2 and hivbase=0;
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=2 and hivbase=1;
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=3 and hivbase=0;
run;

proc univariate data=cleanthesis;
var firstsex;
where htidgroup=3 and hivbase=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var firstsex;
```

```
run;

proc ttest data=cleanthesis;
class hivbase;
var firstsex;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var firstsex;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class hivbase;
var firstsex;
where htidgroup=3;
run;

*Number of lifetime partners;
*By sex;
proc ttest data=cleanthesis;
class sex;
var lifesex;
run;

proc ttest data=cleanthesis;
class sex;
var lifesex;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class sex;
var lifesex;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class sex;
var lifesex;
where htidgroup=3;
run;

proc univariate data=cleanthesis;
class sex;
var lifesex;
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=1 and sex='M';
run;

proc univariate data=cleanthesis;
var lifesex;
```

```
where htidgroup=1 and sex='F';
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=2 and sex='M';
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=2 and sex='F';
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=3 and sex='M';
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=3 and sex='F';
run;

*By HIV status;
proc univariate data=cleanthesis;
class hivbase;
var lifesex;
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=1 and hivbase=0;
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=1 and hivbase=1;
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=2 and hivbase=0;
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=2 and hivbase=1;
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=3 and hivbase=0;
run;

proc univariate data=cleanthesis;
var lifesex;
where htidgroup=3 and hivbase=1;
```

```
run;

proc ttest data=cleanthesis;
class hivbase;
var lifesex;
run;

proc ttest data=cleanthesis;
class hivbase;
var lifesex;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var lifesex;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class hivbase;
var lifesex;
where htidgroup=3;
run;

*Number of partners in the past year;
*By sex;
proc ttest data=cleanthesis;
class sex;
var yearsex;
run;

proc ttest data=cleanthesis;
class sex;
var yearsex;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class sex;
var yearsex;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class sex;
var yearsex;
where htidgroup=3;
run;

proc univariate data=cleanthesis;
class sex;
var yearsex;
run;

proc univariate data=cleanthesis;
var yearsex;
```

```
where htidgroup=1 and sex='M';
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=1 and sex='F';
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=2 and sex='M';
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=2 and sex='F';
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=3 and sex='M';
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=3 and sex='F';
run;

*By HIV status;
proc univariate data=cleanthesis;
class hivbase;
var yearsex;
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=1 and hivbase=0;
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=1 and hivbase=1;
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=2 and hivbase=0;
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=2 and hivbase=1;
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=3 and hivbase=0;
```

```
run;

proc univariate data=cleanthesis;
var yearsex;
where htidgroup=3 and hivbase=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var yearsex;
run;

proc ttest data=cleanthesis;
class hivbase;
var yearsex;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var yearsex;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class hivbase;
var yearsex;
where htidgroup=3;
run;

*Number of partners in the past 3 months;
*By sex;
proc ttest data=cleanthesis;
class sex;
var sex3mo;
run;

proc ttest data=cleanthesis;
class sex;
var sex3mo;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class sex;
var sex3mo;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class sex;
var sex3mo;
where htidgroup=3;
run;

proc univariate data=cleanthesis;
class sex;
```

```
var sex3mo;
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=1 and sex='M';
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=1 and sex='F';
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=2 and sex='M';
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=2 and sex='F';
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=3 and sex='M';
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=3 and sex='F';
run;

*By HIV status;
proc univariate data=cleanthesis;
class hivbase;
var sex3mo;
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=1 and hivbase=0;
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=1 and hivbase=1;
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=2 and hivbase=0;
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=2 and hivbase=1;
```

```
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=3 and hivbase=0;
run;

proc univariate data=cleanthesis;
var sex3mo;
where htidgroup=3 and hivbase=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var sex3mo;
run;

proc ttest data=cleanthesis;
class hivbase;
var sex3mo;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var sex3mo;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class hivbase;
var sex3mo;
where htidgroup=3;
run;

*Sexual behavior variables;
proc sort data=cleanthesis;
by sex;
run;

*All groups;
*By sex;
proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
by sex;
run;

*ttest;
proc ttest data=cleanthesis;
class sex;
var othersex othersexnc othersexwc nocondomsex condomsex;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=1;
```

```
run;

*ttest;
proc ttest data=cleanthesis;
class sex;
var othersex othersexnc othersexwc nocondomsex condomsex;
where htidgroup=1;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=2;
run;

*ttest;
proc ttest data=cleanthesis;
class sex;
var othersex othersexnc othersexwc nocondomsex condomsex;
where htidgroup=2;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=3;
run;

*ttest;
proc ttest data=cleanthesis;
class sex;
var othersex othersexnc othersexwc nocondomsex condomsex;
where htidgroup=3;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=3;
by sex;
run;

*All groups;
*By HIV status;
proc sort data=cleanthesis;
by hivbase;
```

```
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
by hivbase;
run;

*ttest;
proc ttest data=cleanthesis;
class hivbase;
var othersex othersexnc othersexwc nocondomsex condomsex;
run;

*htidgroup=1;
*ttest;
proc ttest data=cleanthesis;
class hivbase;
var othersex othersexnc othersexwc nocondomsex condomsex;
where htidgroup=1;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
*ttest;
proc ttest data=cleanthesis;
class hivbase;
var othersex othersexnc othersexwc nocondomsex condomsex;
where htidgroup=2;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
*ttest;
proc ttest data=cleanthesis;
class hivbase;
var othersex othersexnc othersexwc nocondomsex condomsex;
where htidgroup=3;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where htidgroup=3;
by hivbase;
run;
```

```
*PAST MEDICAL HISTORY;

*Gonorrhea report;
*All groups;
proc freq data=cleanthesis;
tables gonoreport/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables gonoreport/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables gonoreport/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=2;
```

```
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables gonoreport/missing;
where htidgroup=3;
by hivbase;
run;

*Syphilis report;
*All groups;
proc freq data=cleanthesis;
tables syphreport/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables syphreport/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables syphreport/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables syphreport/missing;
```

```
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables syphreport/missing;
where htidgroup=3;
by hivbase;
run;

*Ulcer report;
*All groups;
proc freq data=cleanthesis;
tables ulcerreport/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables ulcerreport/missing;
by sex;
run;

*All groups by HIV status;
```

```
proc freq data=cleanthesis;
tables ulcerreport/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=3;
by sex;
run;
```

```
*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables ulcerreport/missing;
where htidgroup=3;
by hivbase;
run;

*Discharge report;
*All groups;
proc freq data=cleanthesis;
tables discreport/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables discreport/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables discreport/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=2;
by sex;
```

```
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables discreport/missing;
where htidgroup=3;
by hivbase;
run;

*Cystitis/Dysuria report;
*All groups;
proc freq data=cleanthesis;
tables cysdreport/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables cysdreport/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables cysdreport/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=1;
```

```
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables cysdreport/missing;
where htidgroup=3;
by hivbase;
run;

*Lower abdominal pain report;
*All groups;
proc freq data=cleanthesis;
tables lapainreport/missing;
run;

*All groups by sex;
```

```
proc freq data=cleanthesis;
tables lapainreport/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables lapainreport/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=3;
run;
```

```
*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables lapainreport/missing;
where htidgroup=3;
by hivbase;
run;

*ttest;
*By sex, all groups;
proc ttest data=cleanthesis;
class sex;
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
run;

*By HIV status, all groups;
proc ttest data=cleanthesis;
class hivbase;
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
run;

*By sex, htidgroup=1;
proc ttest data=cleanthesis;
class sex;
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
where htidgroup=1;
run;

*By HIV status htidgroup=1;
proc ttest data=cleanthesis;
class hivbase;
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
where htidgroup=1;
run;

*ttest;
*By sex, htidgroup=2;
proc ttest data=cleanthesis;
class sex;
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
where htidgroup=2;
run;

*By HIV status htidgroup=2;
proc ttest data=cleanthesis;
class hivbase;
```

```
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
where htidgroup=2;
run;

*ttest;
*By sex, htidgroup=3;
proc ttest data=cleanthesis;
class sex;
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
where htidgroup=3;
run;

*By HIV status htidgroup=3;
proc ttest data=cleanthesis;
class hivbase;
var gonoreport syphreport ulcerreport discreport cysdreport
lapainreport;
where htidgroup=3;
run;

*PHYSICAL EXAM;

*Inflammation;
*All groups;
proc freq data=cleanthesis;
tables infcomp/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables infcomp/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables infcomp/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
```

```
tables infcomp/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables infcomp/missing;
where htidgroup=3;
by hivbase;
run;

*Ulceration;
*All groups;
proc freq data=cleanthesis;
tables ulcercomp/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables ulcercomp/missing;
by sex;
run;
```

```
*All groups by HIV status;
proc freq data=cleanthesis;
tables ulcercomp/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=3;
by sex;
```

```
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables ulcercomp/missing;
where htidgroup=3;
by hivbase;
run;

*Discharge;
*All groups;
proc freq data=cleanthesis;
tables discomp/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables discomp/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables discomp/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=2;
```

```
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables discomp/missing;
where htidgroup=3;
by hivbase;
run;

*Inflammation composite;
*All groups;
proc freq data=cleanthesis;
tables pgeninf/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables pgeninf/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables pgeninf/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables pgeninf/missing;
```

```
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables pgeninf/missing;
where htidgroup=3;
by hivbase;
run;

*Ulceration composite;
*All groups;
proc freq data=cleanthesis;
tables pgnulc/missing;
run;
```

```
*All groups by sex;
proc freq data=cleanthesis;
tables pgnulc/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables pgnulc/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=3;
run;
```

```
*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables pgnulc/missing;
where htidgroup=3;
by hivbase;
run;

*By sex;
*ttest;
proc ttest data=cleanthesis;
class sex;
var infcomp ulcercomp discomp pgeninf pgnulc;
run;

proc ttest data=cleanthesis;
class sex;
var infcomp ulcercomp discomp pgeninf pgnulc;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class sex;
var infcomp ulcercomp discomp pgeninf pgnulc;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class sex;
var infcomp ulcercomp discomp pgeninf pgnulc;
where htidgroup=3;
run;

*By HIV status;
*ttest;
proc ttest data=cleanthesis;
class hivbase;
var infcomp ulcercomp discomp pgeninf pgnulc;
run;

proc ttest data=cleanthesis;
class hivbase;
var infcomp ulcercomp discomp pgeninf pgnulc;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var infcomp ulcercomp discomp pgeninf pgnulc;
where htidgroup=2;
```

```
run;

proc ttest data=cleanthesis;
class hivbase;
var infcomp ulcercomp discomp pgeninf pgnulc;
where htidgroup=3;
run;

*LABORATORY;

*Gonorrhea lab test;
*All groups;
proc freq data=cleanthesis;
tables gontest/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables gontest/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables gontest/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables gontest/missing;
```

```
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables gontest/missing;
where htidgroup=3;
by hivbase;
run;

*Syphilis (rpr) lab test;
*All groups;
proc freq data=cleanthesis;
tables rpr/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables rpr/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables rpr/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
```

```
tables rpr/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables rpr/missing;
where htidgroup=3;
by hivbase;
run;

*HSV lab test;
*All groups;
proc freq data=cleanthesis;
tables hsv/missing;
run;
```

```
*All groups by sex;
proc freq data=cleanthesis;
tables hsv/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables hsv/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=3;
```

```
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables hsv/missing;
where htidgroup=3;
by hivbase;
run;

*Candida lab test;
*All groups;
proc freq data=cleanthesis;
tables candida/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables candida/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables candida/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables candida/missing;
```

```
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables candida/missing;
where htidgroup=3;
by hivbase;
run;

*Bacterial vaginosis lab test;
*All groups;
proc freq data=cleanthesis;
tables bacvag/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables bacvag/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables bacvag/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
```

```
tables bacvag/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=2;
by hivbase;
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables bacvag/missing;
where htidgroup=3;
by hivbase;
run;
```

```
*Trich lab test;
*All groups;
proc freq data=cleanthesis;
tables trichtest/missing;
run;

*All groups by sex;
proc freq data=cleanthesis;
tables trichtest/missing;
by sex;
run;

*All groups by HIV status;
proc freq data=cleanthesis;
tables trichtest/missing;
by hivbase;
run;

*htidgroup=1;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=1;
run;

*htidgroup=1 by sex;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=1;
by sex;
run;

*htidgroup=1 by HIV status;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=1;
by hivbase;
run;

*htidgroup=2;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=2;
run;

*htidgroup=2 by sex;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=2;
by sex;
run;

*htidgroup=2 by HIV status;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=2;
by hivbase;
```

```
run;

*htidgroup=3;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=3;
run;

*htidgroup=3 by sex;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=3;
by sex;
run;

*htidgroup=3 by HIV status;
proc freq data=cleanthesis;
tables trichtest/missing;
where htidgroup=3;
by hivbase;
run;

*By sex;
*ttest;
proc ttest data=cleanthesis;
class sex;
var gontest rpr hsv candida bacvag trichtest;
run;

proc ttest data=cleanthesis;
class sex;
var gontest rpr hsv candida bacvag trichtest;
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class sex;
var gontest rpr hsv candida bacvag trichtest;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class sex;
var gontest rpr hsv candida bacvag trichtest;
where htidgroup=3;
run;

*By HIV status;
*ttest;
proc ttest data=cleanthesis;
class hivbase;
var gontest rpr hsv candida bacvag trichtest;
run;

proc ttest data=cleanthesis;
class hivbase;
var gontest rpr hsv candida bacvag trichtest;
```

```
where htidgroup=1;
run;

proc ttest data=cleanthesis;
class hivbase;
var gontest rpr hsv candida bacvag trichtest;
where htidgroup=2;
run;

proc ttest data=cleanthesis;
class hivbase;
var gontest rpr hsv candida bacvag trichtest;
where htidgroup=3;
run;

*Frequencies and percents for table 2;
*SEXUAL BEHAVIOR;
proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where hivbase=0;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where hivbase=1;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where hivbase=0 and seroconvert=0;
run;

proc freq data=cleanthesis;
tables othersex othersexnc othersexwc nocondomsex condomsex/missing;
where hivbase=0 and seroconvert=1;
run;

*PAST MEDICAL HISTORY;
proc freq data=cleanthesis;
tables gonoreport syphreport ulcerreport discreport cysdreport
lapainreport/missing;
where hivbase=0;
run;

proc freq data=cleanthesis;
tables gonoreport syphreport ulcerreport discreport cysdreport
lapainreport/missing;
where hivbase=1;
run;

*PHYSICAL EXAM;
proc freq data=cleanthesis;
tables infcomp ulcercomp discomp pgeninf pgnulc/missing;
where hivbase=0;
run;

proc freq data=cleanthesis;
```

```
tables infcomp ulcercomp discomp pgeninf pgnulc/missing;
where hivbase=1;
run;

*LABORATORY;
proc freq data=cleanthesis;
tables gontest rpr hsv candida bacvag trichtest/missing;
where hivbase=0;
run;

proc freq data=cleanthesis;
tables gontest rpr hsv candida bacvag trichtest/missing;
where hivbase=1;
run;

************************************;
*****          MODEL            *****;
************************************;

*********************************************************************
*******;
*1) Genital abnormalities and STIs as predictors of HIV prevalence at
baseline;
*********************************************************************
*******;

*UNIVARIATE ODDS RATIOS AND CI's for table 3;
*PAST MEDICAL HISTORY;
*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=gonoreport;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=gonoreport;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=syphreport;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=syphreport;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=ulcerreport;
```

```
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=ulcerreport;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=discreport;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=discreport;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=cysdreport;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=cysdreport;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=lapainreport;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=lapainreport;
by htidgroup;
strata sexnum;
run;

*PHYSICAL EXAM;
*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=infcomp;
strata sexnum;
run;

*By htidgroup;
```

```
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=infcomp;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=ulcercomp;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=ulcercomp;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=discomp;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=discomp;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=pgeninf;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=pgeninf;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=pgnulc;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=pgnulc;
by htidgroup;
strata sexnum;
run;
```

```
*LABORATORY;
*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=gontest;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=gontest;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=rpr;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=rpr;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=hsv;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=hsv;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=candida;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=candida;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=bacvag;
```

```
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=bacvag;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=trichtest;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=trichtest;
by htidgroup;
strata sexnum;
run;

*SEXUAL BEHAVIOR;
*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=othersex;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=othersex;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=othersexnc;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=othersexnc;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=othersexwc;
strata sexnum;
run;

*By htidgroup;
```

```
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=othersexwc;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=nocondomsex;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=nocondomsex;
by htidgroup;
strata sexnum;
run;

*All groups;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=condomsex;
strata sexnum;
run;

*By htidgroup;
proc logistic data=cleanthesis covout outest=info;
model hivbase (event='1')=condomsex;
by htidgroup;
strata sexnum;
run;

*Initial model: predictors of HIV prevalence at baseline among all
groups
 Variables to use determined by significant results in univariate
analysis;
proc logistic data=cleanthesis covout outest=info;
class
gonoreport (param=ref ref='0')
syphreport (param=ref ref='0')
ulcerreport (param=ref ref='0')
discreport (param=ref ref='0')
cysdreport (param=ref ref='0')
lapainreport (param=ref ref='0')
infcomp (param=ref ref='0')
ulcercomp (param=ref ref='0')
pgnulc (param=ref ref='0')
rpr (param=ref ref='0')
trichtest (param=ref ref='0');
model HIVbase (event='1')=gonoreport syphreport ulcerreport discreport
cysdreport
      lapainreport infcomp ulcercomp pgnulc rpr trichtest/rl;
strata sexnum;
run;
```

```
*Initial model: predictors of HIV prevalence at baseline, HTID group=1
 Variables to use determined by significant results in univariate
analysis;
proc logistic data=cleanthesis covout outest=info;
class
gonoreport (param=ref ref='0')
syphreport (param=ref ref='0')
ulcerreport (param=ref ref='0')
cysdreport (param=ref ref='0')
lapainreport (param=ref ref='0')
infcomp (param=ref ref='0')
ulcercomp (param=ref ref='0')
pgnulc (param=ref ref='0')
rpr (param=ref ref='0')
trichtest (param=ref ref='0');
model HIVbase (event='1')=gonoreport syphreport ulcerreport cysdreport
      lapainreport infcomp ulcercomp pgnulc rpr trichtest/rl;
strata sexnum;
where htidgroup=1;
run;

*Initial model: predictors of HIV prevalence at baseline, HTID group=2
 Variables to use determined by significant results in univariate
analysis;
proc logistic data=cleanthesis covout outest=info;
class
syphreport (param=ref ref='0')
ulcerreport (param=ref ref='0')
discreport (param=ref ref='0')
cysdreport (param=ref ref='0')
infcomp (param=ref ref='0')
ulcercomp (param=ref ref='0')
pgnulc (param=ref ref='0')
rpr (param=ref ref='0')
hsv (param=ref ref='0');
model HIVbase (event='1')=syphreport ulcerreport discreport cysdreport
      infcomp ulcercomp pgnulc rpr hsv/rl;
strata sexnum;
where htidgroup=2;
run;

*Initial model: predictors of HIV prevalence at baseline, HTID group=3
 Variables to use determined by significant results in univariate
analysis;
proc logistic data=cleanthesis covout outest=info;
class
ulcerreport (param=ref ref='0')
ulcercomp (param=ref ref='0')
pgnulc (param=ref ref='0')
othersex (param=ref ref='0');
model HIVbase (event='1')=ulcerreport ulcercomp pgnulc othersex/rl;
strata sexnum;
where htidgroup=3;
run;
```

```
*Checking collinearity: initial, full model among all groups;
%include 'H:\Thesis\collingenmodv9c.sas';

proc logistic data=cleanthesis covout outest=info;
class
gonoreport (param=ref ref='0')
syphreport (param=ref ref='0')
ulcerreport (param=ref ref='0')
discreport (param=ref ref='0')
cysdreport (param=ref ref='0')
lapainreport (param=ref ref='0')
infcomp (param=ref ref='0')
ulcercomp (param=ref ref='0')
pgnulc (param=ref ref='0')
rpr (param=ref ref='0')
trichtest (param=ref ref='0');
model HIVbase (event='1')=gonoreport syphreport ulcerreport discreport
cysdreport
      lapainreport infcomp ulcercomp pgnulc rpr trichtest/rl;
strata sexnum;
run;

%collin (covdsn=info);
run;
*None of the CNI's are above 30 - No collinearity;
title;

*Initial model: predictors of HIV prevalence at baseline among all
groups, stepwise selection;
proc logistic data=cleanthesis covout outest=info;
class
gonoreport (param=ref ref='0')
syphreport (param=ref ref='0')
ulcerreport (param=ref ref='0')
discreport (param=ref ref='0')
cysdreport (param=ref ref='0')
lapainreport (param=ref ref='0')
infcomp (param=ref ref='0')
ulcercomp (param=ref ref='0')
pgnulc (param=ref ref='0')
rpr (param=ref ref='0')
trichtest (param=ref ref='0');
model HIVbase (event='1')=gonoreport syphreport ulcerreport discreport
cysdreport
      lapainreport infcomp ulcercomp pgnulc rpr trichtest/rl
selection=stepwise;
strata sexnum;
run;
*Keep: gonoreport, ulcerreport, cysdreport, rpr;

*Checking collinearity: reduced, all groups model;
%include 'H:\Thesis\collingenmodv9c.sas';

proc logistic data=cleanthesis covout outest=info;
class
gonoreport (param=ref ref='0')
ulcerreport (param=ref ref='0')
```

```
cysdreport (param=ref ref='0')
rpr (param=ref ref='0');
model HIVbase(event='1')=gonoreport ulcerreport cysdreport rpr /rl;
run;

%collin (covdsn=info);
run;
*None of the CNI's are above 30 - No collinearity;
title;

*Final model: predictors of HIV prevalence at baseline among all
groups;
proc logistic data=cleanthesis covout outest=info;
class
gonoreport (param=ref ref='0')
ulcerreport (param=ref ref='0')
cysdreport (param=ref ref='0')
rpr (param=ref ref='0');
model HIVbase(event='1')=gonoreport ulcerreport cysdreport rpr /rl;
run;
```