Factors Associated with Bacterial Vaginosis in Sexually Active Women Enrolled in the Microbial Development Program 301 Study

by

Mercy Manyema

Student Number: 469168

A research report submitted in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE
IN
INFECTIOUS DISEASE EPIDEMIOLOGY

SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF THE WITWATERSRAND
SEPTEMBER 2013

University of the Witwatersrand, Faculty of Health Sciences,

School of Public Health

7 York Road, Johannesburg

September 2013

DECLARATION

I, Mercy Manyema (student number 469168) am submitting my research report in partial fulfilment of the requirements of the MSc in the field of Infectious Disease Epidemiology at the University Of Witwatersrand School Of Public Health. This report has not been submitted before for any degree at any other university. I hereby declare that this research report is my own work. Where I have used the thoughts or ideas of others, the required referencing conventions have been adhered to.

Signed: 

Date: 26 September 2013

DEDICATION

I dedicate this work to my family: my husband Shaw Manyema and my two children Tatenda and Tinashe Manyema for the sacrifices, support and encouragement they gave me as I studied. This work is also dedicated to all those working tirelessly to improve the sexual reproductive health of women.

Mercy Manyema

September 2013

ABSTRACT

Introduction

Bacterial vaginosis (BV) is a highly prevalent vaginal infection which poses a significant public health burden in Sub-Saharan Africa (SSA) due to its association with HIV, other STIs and several gynaecological and obstetrical complications. The aim of this study was to explore the underlying and proximate factors associated with BV and the relationships between them.

Materials and Methods

This study was a cross-sectional secondary analysis of the data collected during the Microbial Development Program (MDP) 301 trial. Logistic regression and structural equation modelling were used to test for the associations between BV and the explanatory variables and to test for the direct, indirect and total effects of the variables on BV.

Results

A total of 2 470 women were included in the analysis and of these 2 203 were aged 40 and below. The majority of them were unemployed at 72% and 51,8% were in the lowest socio-economic level. The baseline prevalence of BV was 40.5%. In the logistic regression, high socio-economic level (AOR=1.66; 95% CI 1.04-2.64) and using a condom during their last sexual encounter (AOR 0.82; 95% CI 0.69-0.97) were associated with BV infection. The STIs significantly associated with BV infection were: Herpes Simplex Virus 2 (HSV2) (AOR=1.31; 95% CI 1.10-1.56), trichomoniasis (AOR=2.68; 95% CI 1.97-3.64) and chlamydia infection (AOR 2.02; 95% CI 1.61-2.62). In the structural equation modelling (SEM) high socio-economic status had a positive direct effect on BV infection (beta=0.12, OR=1.14).

Condom use during the last sex act had a negative direct effect on BV (beta=-0.043, OR=0.96). The presence of *T.vaginalis*, HSV2 or chlamydia infection had significant positive effects on BV infection.

Conclusions

Sexual behavioural factors and the presence of STIs were significantly associated with BV. The SEM analysis showed the interaction of contraceptive use and sexual behavioural factors. No interaction between the STIs and sexual behaviour could be demonstrated in this study.

Recommendations

Further study is necessary to establish temporal relationships between the variables found to impact each other in the SEM analysis.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr Eustasius Musenge and Dr Harry Moultrie for their guidance and input in the conducting of the research and the preparation of this report. The lecturers at the Wits School of Public Health played an important role in laying the theoretical foundation and providing much needed insight, and for that I am grateful. Many thanks go to the administrative staff for their assistance throughout the duration of the course.

I also would like to acknowledge and extend my gratitude to the Wits Reproductive Health and HIV Institute and the MDP Trial 301 for the data used in this study.

Last but not least, I am grateful to God Almighty for His mercy and provision.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AOR	Adjusted odds ratio
BV	Bacterial vaginosis
HIV	Human immunodeficiency virus
MDP	Microbicide Development Program
OR	Odds ratio
SEM	Structural equation modelling
SEMs	Structural equation models
SSA	Sub-Saharan Africa
STI	Sexually transmitted infections
HSV2	Herpes simplex virus 2
T. vaginalis	Trichomonas vaginalis
N. gonorrhoeae	Neisseria gonorrhoeae
SES	Socio-economic status

DEFINITION OF TERMS

Proximate determinant/variables	Also called intermediate variables, are variables through which underlying variables operate to influence biological processes that cause disease
Underlying variables	Also called contextual variables, are socio-demographic and cultural variables that contribute to the occurrence of disease
Observed variable	A variable that is in the data set
Latent variable	A variable not in the data set i.e. unobserved
Exogenous variable	A variable from which paths only originate and no paths point to it
Endogenous variable	A variable to which any path points

CHAPTER 1: INTRODUCTION

Introduction

In this chapter, the burden and public health implications of bacterial vaginosis (BV) world-wide, in Sub-Saharan Africa (SSA) and in South Africa are highlighted. Published literature on the factors associated with BV is reviewed. The relevance and usefulness of a conceptual framework in analysing the factors associated with BV is proposed. Structural equation modelling is introduced as a tool to define the relationships between these factors and the chapter concludes with a description of the objectives of the study.

1.1 Background

Bacterial vaginosis is one of the commonest causes of vaginal infection and abnormal vaginal discharge in women of child-bearing age (1-3). Predominantly affecting young, sexually active women, its prevalence varies with the population of women being studied e.g. sex workers, HIV-positive women, pregnant women or students (1, 4-6). A population-based study in rural Uganda found a prevalence of 50.9% (7). In sexually transmitted disease (STD) clinic attendees, the prevalence of BV has been found to be between 24-40%, and 16-29% in pregnant women (4). A South African study found a baseline prevalence of 41.1% in HIV-positive women (8).

BV is characterised by a reduction in the concentration of hydrogen peroxide-producing lactobacilli and a corresponding increase in the concentration of *Gardnerella vaginalis; Mycoplasma hominis; Ureaplasma urealyticum;* anaerobic rods of the *Prevotella, Mobiluncus* and *Bacteriodes species;* and anaerobic *Peptostreptococcus* species (4). Despite advances in the knowledge of its pathosphysiology, the aetiology of BV continues to be poorly understood (4, 9).

1.1 Statement of the Problem

The prevalence of BV is high in SSA. A prospective study jointly conducted in Zimbabwe and Uganda in a population of both HIV-negative and positive women showed a prevalence of between 19-29% (10). In a study in Kenya the prevalence was found to be 44% (6). The prevalence of BV among female sex workers was found to be even higher at 62.9% in a Tanzanian study (1). In South Africa, prevalence rates of between 41 and 74% in HIV positive women have been reported (11-13).

An increasing body of knowledge points to the association between BV and sexually transmitted infections (STIs), including HIV. An increased risk of HIV sero-conversion among women with BV has been found, as well as a greater than three-fold increased risk of female to male HIV-1 transmission (8, 14). An association has also been found between BV and incident STIs and upper genital tract infections (15, 16). Furthermore, BV has been linked with many adverse obstetrical and gynaecological outcomes such as premature rapture of the membranes, chorioamnionitis, preterm delivery, low birth weight and pelvic inflammatory disease (4, 6, 17, 18). With the prevalence of BV being high in SSA, the sequelae of this often asymptomatic infection pose a serious public health burden.

1.2 Justification

The prevention and treatment of BV in South Africa depends on the knowledge of its risk factors, and the factors associated with it, especially those that can be targeted using interventions that are practicable in resource-poor settings (6, 17). Epidemiological studies that examine the distribution and risk factors of various conditions, including BV, often include large numbers of variables to identify the multiple inter-relationships (19), and multivariable analyses are employed to assess these relationships (20). A conceptual framework is a useful tool to guide the collection, analysis and interpretation of epidemiological data to avoid confusion over the relevance and relative importance of variables (19, 20). A proximate determinants framework provides a logical strategy for handling the complex inter-relationships between disease determinants (21).

Literature is replete with evidence of the factors associated with BV but silent on the use of a conceptual framework to analyse these factors. This study aims to examine the factors associated with BV and to understand the relationship between the underlying and proximate factors. Structural equation models will be used as an additional tool to logistic regression to explore and analyse these relationships. This information can inform the formulation of interventions directed at these factors in the prevention of BV and its associated direct and indirect (related to HIV and other STIs) complications.

1.3 Literature Review

The aetiology, pathogenesis and diagnosis of bacterial vaginosis has been extensively studied and presented in literature. This literature will be reviewed,

together with the factors commonly found to be associated with bacterial vaginosis.

1.3.1 Bacterial Vaginosis

1.3.1.1 Aetiology and Pathogenesis

Bacterial vaginosis is a polymicrobial clinical condition in which there is an overgrowth of anaerobic bacteria, accompanied by a reduction in the normal *Lactobacillus* population (1, 2, 4, 9, 22). An increased production of proteolytic carboxylase enzymes leads to the breakdown of vaginal peptides to several amines which are volatile and malodorous in high pH (4). The unpleasant “fishy-smelling” discharge that most women present with results from amine-induced vaginal transudation and squamous epithelial cell exfoliation.

It remains unknown, however, whether the reduction in numbers of *Lactobacillus* precedes or follows the complex change in microflora, and which of the organisms present are key to the development of BV (4, 22).

1.3.1.2 Diagnosis

Most women with BV present with a thin, homogenous off-white discharge, but about 35% of the infections are asymptomatic (4, 23). Diagnosis of the condition includes clinical examination and several laboratory diagnostic criteria. The Amsel criteria have been the most commonly used method to diagnose BV and are based on a wet mount of the vaginal discharge. The composite criteria include a thin homogenous discharge, elevated vaginal pH above 4.5, a fishy smell on addition of potassium hydroxide and the presence of clue cells in the wet mount (24). Three of these criteria need to be present for a positive diagnosis of BV to be made.

The Nugent and Hay-Ison criteria are based on a Gram-stained smear which is graded according to the bacteria present and their morphology (4, 24, 25). The Hay-Ison criteria were used in this study to define BV.

1.3.2 Factors Associated with BV

Extensive research has been done on the factors associated with or that increase the risk of BV infection. Studies have been done in developing as well as developed countries (1, 3, 5, 6). Several different populations of women have been studied including sex workers, rural women, HIV positive and HIV negative women (1, 8, 14, 26). A number of factors have been shown to be associated with bacterial vaginosis, with most studies finding multiple and diverse factors.

1.3.2.1 Demographic and Socio-economic Factors

The National Health and Nutrition Examination Survey (NHNES) conducted in America between 2001 and 2004 revealed that the prevalence of BV varied with age, and education and poverty levels (5). Fewer years of education and being of black ethnic race were identified by other studies as being independently associated with BV (9, 23). The studies also revealed that the association of BV with ethnicity was more significant as the number of years of education decreased pointing to the presence of other unknown mediating factors of this association (23). It is also postulated that inter-racial and inter-country differences in BV infection arise due to culturally-based vaginal douching practices, with women of black origin more likely to douche than white ethnic women (27, 28).

1.3.2.2 Sexual Behaviour, Condom Use and Contraceptive Use Factors

Although the classification of BV as a sexually transmitted infection still remains a contentious issue, BV has been found to be associated with sexual behavioural

factors. According to a study done in Australia, high risk sexual behaviour, as indicated by new sexual partner and increased numbers of partners in the past year, and a high number of lifetime sexual partners, is significantly associated with BV infection (3).

Similar results were found in a cross-sectional study done in Sweden (29). A short-term relationship before and after sexual debut, multiple partners in the past month and a high number of lifetime sexual partners were all associated with BV (29). A systematic review and meta-analysis of the association between sexual risk factors and bacterial vaginosis revealed a strong positive association between history of male sexual partners and BV, and history of female sexual partners and BV (18). Condom-use was found to be protective against BV infection (18).

Receptive anal intercourse before vaginal intercourse as well as sex with an uncircumcised male partner have been found to be independently associated with acquisition of BV (9).

Hormonal or oral contraceptives have largely been found to be protective against BV while intrauterine devices increase the risk of infection or have no effect (23, 30, 31).

1.3.2.3 Vaginal Cleansing Practices

Vaginal cleansing practices like vaginal douching are also associated with BV. In a study done in HIV positive women in Nigeria, vaginal douching was linked to increased BV risk (2). Douching in the past two and six months were significant predictors of BV infection in the NHNES study done in the USA and a study in Michigan respectively (5, 23). Increased frequency of BV infection was linked to douching in the last month in another study, with those that had douched in the

past week having the highest risk (32). A study in Tanzanian female facility workers however found no association between BV and vaginal cleansing (1), as well as another carried out in young Thai women (17).

It is thought that vaginal cleansing practices introduce inter-racial and inter-country differences in the prevalence of BV because black ethnic women are more likely to douche than their white counterparts (28).

1.3.2.4 Biological Factors

Though the interactions are not well understood, the presence of sexually transmitted infections (STIs) has been shown to be associated with BV in several studies. For example, a study done in female facility workers found that *Trichomonas vaginalis* (*T. vaginalis*) infection was an independent risk factor for BV (1). Investigators in Alabama, USA, found that women with prevalent *T. vaginalis* infection were 2.8 times more likely to have BV than those who did not and that those with incident syphilis were 9.7 times more likely to have BV than those who did not (33). Patients with STIs were more prone to have vaginal flora indicative of BV as found in another study conducted in Japanese women (34). The BV rate was high in women with *Chlamydia trachomatis (C. trachomatis), Neisseria gonorrhoeae* (*N. gonorrhoeae)* and *T. vaginalis* infections(34). In Kenya the odds of having BV was 13.2 times higher if the woman had trichomoniasis (6).

1.3.2.5 Other Factors

A study in young Thai women found intercourse during menstruation, male partners having sex with other women, and smoking to be independently associated with BV (17).

1.3.3 Proximate determinants framework

First postulated by Henry Mosley and Lincoln Chen as an analytical framework for studying determinants of child mortality (35), the proximate determinants framework is increasingly becoming popular in studying risk factors of infectious diseases (19, 20). It has been widely used in the study of fertility and child survival in developing countries (20).

The analytical framework as proposed by Mosley and Chen integrates medical and social science methods, and incorporates both social and biological variables in the analysis (35).

The underlying premise of the framework is that all social and economic determinants operate through a common set of 'proximate' or intermediate variables to impact the outcome (child survival in their case) (35).This framework has been adopted and adapted for use in the study of sexually transmitted infections, particularly HIV infection as the importance of the role of underlying socioeconomic and cultural determinants and other contextual variables in the transmission of HIV and other sexually transmitted infections (STIs) is becoming more apparent (19, 20, 36). The proximate determinants link the underlying variables to the actual biological factors that lead to the process of the spread of infection (19).

The proximate determinants framework as postulated by Mosley and Chen, and Boerma was used in this study to describe how underlying variables are associated with BV infection and in assessing relationships between various underlying and proximate factors (20, 35).

The framework was used to define the structural equation model and inform which variables from the data were to be treated as underlying, and which were to be treated as proximate.

1.4.1 Structural equation models

1.4.1.1 Description

Structural equation modelling (SEM) is an important and powerful statistical tool for evaluating complex relations in several research areas (37).
Structural equation models (SEMs) are a flexible class of models that allow complex modelling of multivariate data and multiple, closely related predictors (38). SEM provides a flexible framework for developing and analysing complex relationships among multiple variables simultaneously, while also managing measurement error, which is one of the greatest limitations of most studies (39).

SEM was developed in the early 1900s with the development of factor analysis by Spearman in 1904, and path analysis by Wright in 1918 (40, 41). Advances in computer programming, the increased availability and ease of use of software has resulted in more and more researchers using the technique since then (39).

SEM includes multivariate data analysis techniques that combine aspects from multiple regression models and factor analysis to simultaneously estimate a series of relations of dependency (37, 39). Variables in SEMs may be observed or unobserved/ latent. Variables may also be exogenous (explanatory), with no path pointing to them or endogenous (dependent) with at least one path pointing to them.

SEM can be used to analyse data that are obtained experimentally or observationally and allows researchers to postulate possible causal links between variables using observational data (39). Complex mediated mechanisms can be evaluated by disaggregating the effects of variables on each other (37).

1.4.1.2 Application

The application of SEMs involves several stages, including developing a theoretical conceptual model, specification of the mathematical model, determining the model's identifiability and fit, and evaluating its goodness of fit (37).

A research question is identified and outlined based on theory and/or previous empirical findings (39, 42). A theoretical model is then specified and the specific relationships to be tested are determined a priori by the researcher based on current knowledge and previous work on the subject (37, 39, 42). The theoretical model consists of a systematic set of relations among the variables, some direct and some indirect (39). A full structural equation model comprises a structural model and a measurement model.

The relationships are defined by a series of regression equations with the path coefficients (beta coefficients) measuring the importance of a given path from cause to effect (39). Path diagrams are used to visually illustrate the model and summarise the set of hypotheses (37, 43). The symbols used in the path diagrams represent the assumptions of the model as shown in Figure 1.1.

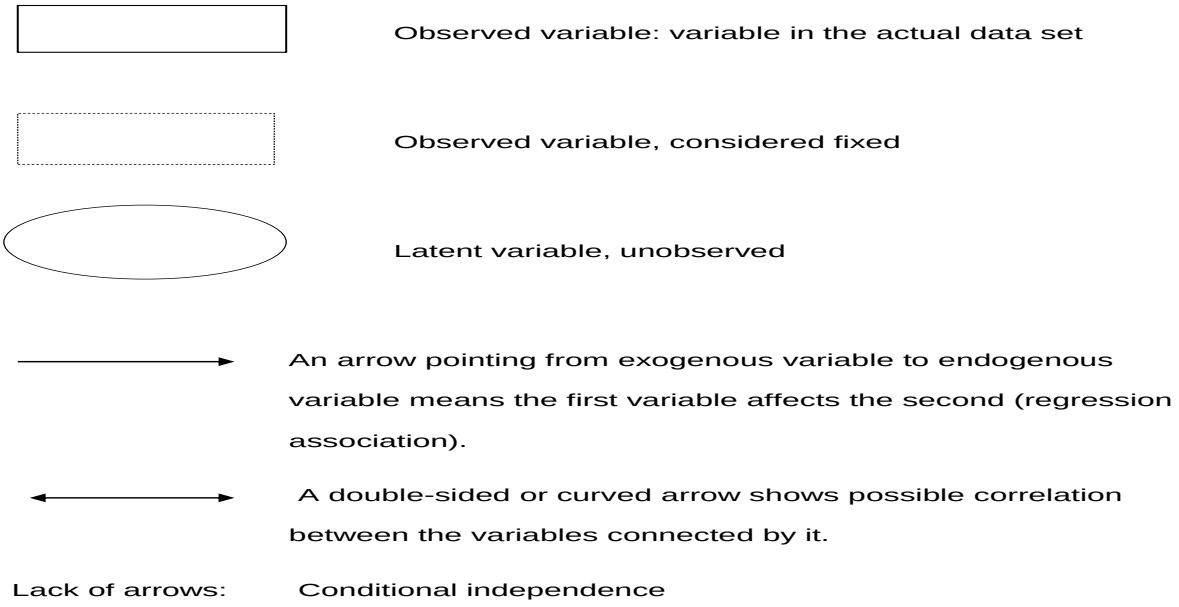


Figure 1.1 Notation Used in Path Diagrams

1.4.2 Conclusion

The prevalence of BV is high in sub-Saharan Africa (1, 6, 8, 14). BV is of significant public health concern due to its association with STIs, including HIV, and other adverse gynaecological and obstetrical conditions (6, 8, 14). Of the several studies that have been carried out in different populations of women to determine the risk factors of, and the factors associated with BV, most of them simply placed the factors in regression models for analysis (2, 6, 17, 23). The proximate determinants framework can be used to guide the analysis of the data to enable the evaluation of the complex inter-relationships between the underlying and proximate determinants of BV. SEM is a powerful analytical set of techniques that can be used to analyse the relationships simultaneously (37, 39). To the best of our knowledge, literature is silent on the use of a conceptual framework or structural equation modelling to analyse factors associated with BV. SEM was used in this study together with logistic regression to analyse these factors.

1.5 Objectives

Aim: To explore the underlying and proximal factors associated with bacterial vaginosis and the relationships between them.

Objectives:

1. **a.** To describe the underlying (demographic and socioeconomic) and proximate (sexual activity, contraceptive use and vaginal cleansing) characteristics of the women in the MDP301 study

   b. To determine the prevalence of BV infection at enrolment
2. To determine the associations between the underlying characteristics and BV infection, and the proximate determinants and BV infection, using logistic regression techniques
3. To investigate the relationships between the underlying and proximate determinants, and their effect on BV infection using structural equation models.

CHAPTER 2: MATERIALS AND METHODS

Introduction

In this chapter the study design, population, sample and study methodology is described. The variables used in the study are described and defined. The methodology of measuring the outcome and explanatory variables is outlined, as well as the data processing and analysis. Ethical considerations conclude the chapter.

2.1 Study design

This study was an analytical cross-sectional study. Secondary analysis of the screening and baseline data collected during the MDP301 clinical trial (44, 45), was performed. This study was a cross-sectional design in that only the baseline data were analysed with no follow-up data being included. The design was appropriate since prevalence ratios could be generated and used in determining key determinants of bacterial vaginosis (BV) infection.

2.2 Study population

The study population included all sexually active women enrolled in the MDP 301 trial in Johannesburg.

2.2.1 MDP 301Trial

The MDP 301 Trial was a multi-centre randomised placebo-controlled Phase 3 trial run by a collaborative network of institutions called the Microbicide Development Programme (MDP) (44).

It was carried out at six study locations across Eastern and Southern Africa, three of which were in South Africa: one in Durban and two in Johannesburg (45). The main objective of the trial was to determine the efficacy and safety of 0.5% and 2% concentrations of the vaginal microbicide gel PRO 2000/5 compared to a placebo in preventing HIV infection (45). PRO2000 gel is a synthetic naphthalene sulphonate polymer with antiviral activity against HIV-1 and STIs (44).

The inclusion criteria for the trial were: 18 years or older at enrolment, sexually active at entry and follow-up, willing to undergo HIV testing at screening and during follow-up, HIV negative at screening according to the local HIV testing algorithm, willing to receive HVI result before randomisation, willing to undergo regular speculum examination, genital infection tests and pregnancy tests, willing to receive health education about condoms and willing and able to give informed consent (45). Excluded from the study were women who: were unable or unwilling to provide a reliable method of contact for follow-up, were likely to permanently move out of study area, were likely to have sex more than 14 times a week on a regular basis during the course of the study, were using spermicides regularly, were pregnant or within 6 weeks post-partum, had grade 3 clinical or laboratory abnormalities, required referral for clinically suspicious cervical lesions or had participated within 30 days in another trial likely to influence the results of this trial.

2.3 Study Sample

The sample for the present study was made up of all sexually active women enrolled in the MDP301 at the two Wits Reproductive Health and HIV Institute (WRHI) sites in Johannesburg between October 2005 and August 2008 (44).

Inclusion criteria:

- Older than 18 years of age
- Had a bacterial vaginosis test done with the Hay-Ison method
- Attended both the screening and enrolment visit

Exclusion criteria:

- Younger than 18 years of age
- Missing BV result
- BV test not done with the Hay-Ison method
- Attended the screening visit but was not enrolled

2.3.1 Power calculations

Power analysis was performed to determine if the study would have sufficient power to detect change in effect size. Both regression and SEM techniques were used in this study therefore the study needed to have sufficient power for both sets of techniques.

The Power Analysis and Sample Size (PASS) Version 11 software was used to determine the power achieved by the study in logistic regression to detect an odds ratio of 1.5 with a sample size of 2500 (46). Assuming the prevalence of BV is 0.45 (1, 6) and that 10% of the sample have the independent variable of interest, 86% power will be achieved in univariate logistic regression analysis.

This software is based on the method by Hsieh (47). Sample size calculations were performed for the MDP 301 Trial and these power calculations were performed after the fact.

It has been proposed that in SEM techniques at least 10 participants be allowed for each variable measured i.e. if there are 20 variables the sample size should be at least 200 (48). Other authors have proposed a 'critical sample size' 200 for SEM analysis (48-50). A sample size of 200 or more is generally accepted as sufficient for SEM.

2.4 Data Management

2.4.1 Measurement

Women in the MDP 301 trial initially attended a screening visit at which eligibility for the trial was assessed, demographic information was collected, a behavioural interview was administered and HIV tests were done (44). A unique trial number was assigned to each participant at this visit.

Women were encouraged to return up to 6 weeks later for the enrolment visit. At the enrolment visit eligibility was re-assessed, a further behavioural interview was administered, a clinical examination was conducted and other STI tests and BV tests were performed (44).

This information was entered onto separate clinical research forms and then into a Structured Query Language (SQL) database using double-entry (44).

The data were transferred to STATA (STATA Corporation, College Station, TX) for analysis, with separate datasets for demographic data, sexual behaviour data and laboratory data. The demographic dataset included data collected at the screening visit while the latter two comprised information collected at both the screening and the enrolment visits. It was these data sets that were received from the WRHI and from which the study variables for this study were extracted. All the data were de-identified and the unique trial numbers used to identify the participants. The data were stored on a password-protected computer during the study.

2.4.2 Cleaning and Extraction of Variables

The sexual behaviour and laboratory datasets were reshaped such that each woman had their screening and enrolment data in a single row, and were then merged together using the unique trial numbers. These datasets were merged with the demographic dataset to form one complete dataset with all the relevant variable information. All women who had missing BV data were removed and excluded from the analysis of this study. One duplicate observation was found and was dropped from the dataset.

2.5 Study Variables

2.5.1 Outcome variable

The outcome variable of interest was BV infection, diagnosed through the Hay-Ison method. The Hay-Ison grading system is based on a Gram-stained smear of a vaginal swab (see Appendix 1). Smears are graded in the following manner (25, 51):

- Grade 0 – epithelial cells with no bacteria seen
- Grade I – lactobacillus morphotype only (normal flora)

- Grade II – reduced lactobacillus morphotype with mixed bacterial morphotypes (intermediate flora)
- Grade III – mixed bacterial morphotypes with few or absent lactobacillus morphotype (BV)
- Grade IV – epithelial cells covered with gram-positive cocci only

For this study all results graded as III were taken to be BV positive, and all results with any other grade were taken as BV negative.

2.5.2 Explanatory Variables

Demographic variables

- **Age** was defined as the age of the woman, in years, at the screening visit. For analysis age was categorised a priori into 10-year intervals due to it being highly skewed to the right. 10-year intervals were used in order to normalise the distribution of age because 5-year categories failed to attenuate the skew.
- **Education** was taken to be the highest level of education attained by the woman at the time of screening. The variable was categorised into three categories: no education/primary; secondary and tertiary/ vocational training.

  Each category included those who had completed the level, and those who had not completed. The categories from the original dataset were collapsed this way to cater for cells with less than 5 observations
- **Number of people sleeping per room** was defined as the number of people sleeping in each room designated for sleeping.

This variable was used as an indicator of over-crowding and was generated by dividing the variable **number of people usually sleeping in household** by **number of rooms used for sleeping.** It was categorised as 1-2 people, and greater than 2 people.

- **Partner's contribution** was defined as the proportion of the household income contributed by the woman's partner. It was categorised into nothing; one third; half; three quarters and all
- **Head of household** was the head of the household where the woman resided. The categories included were self; partner; sibling/child; parent/in-laws and other relative. A category for nephew or niece was collapsed into the child category due to the small numbers.

Socio-economic variables

- **Employment status** was the current type of employment, as described by the women at the time of screening. It was categorised into five categories: full time; part-time; student; work-seeker; and unemployed. One woman declined to answer and was set to missing, and one retired woman was added to the 'Unemployed' category
- **Socio-economic level** was a variable that was used to measure the level of socio-economic status of the woman. A count variable was created using the variables **electricity, radio, television, telephone, refrigerator, personal computer, washing machine, bicycle, motorcycle, a car, a horse/donkey and cattle/sheep.** Three categories were created based on the number of assets and utilities possessed by the woman's household.

The categories were as follows: low - 0 to 3 assets; middle – 4 to 6 assets; high – 7 to 9 assets. None of the women had a total count of assets greater than 9.

Contraceptive use factors

- **Type of family planning method** was a group of variables measuring the number of women using several types of family planning methods. The methods included were **natural rhythm; oral contraceptive pills; foam; the diaphragm; Nur-Isterate injection; Depo Provera injection; other injectables; intra-uterine device; Norplant; condom; traditional oral and vaginal and other traditional methods; sterilisation and other methods**. The variables were binary with 0 representing all the women not using that particular method and 1 representing those using it. Category 0 included all those using other methods or not using one at all. A variable was available in the dataset measuring whether a woman was using at least one method of family planning or not but was omitted from analysis to avoid confounding and collinearity with the variables described above.

Sexual behaviour variables

- **Male sex partner** was a variable that showed whether the woman had a current male sex partner, coded as 0 (no) and 1 (yes).
- **Type of current partner** was defined as the type of male sex partner the woman had: either a long term stable partner or not.
- **Condom use at last sex act** was a binary variable indicating whether the woman had used a condom the last time they had intercourse.

- **Age at sexual debut** was a variable of how old the woman was when she had her first sexual intercourse encounter. It was categorised into <15 years; 15-19 years and 20 years or older.
- **Vaginal insertion** was a binary variable defined as the woman having inserted anything other than water or fingers into their vagina in the past week.
- **Number of times inserted** was defined as the number of times vaginal insertion as defined above occurred. The variable was categorised into five: more than once per day; once per day; less than once per day but more than once per week; once in a week and don't remember.
- **Anal sex** was a binary variable indicating whether or not the woman had engaged in anal intercourse in the past four weeks.
- **Condom use during anal sex** was a variable indicating whether the woman used a condom during anal sex, categorised as always; most of the time; sometimes and never.

2.6 Data Processing and Analysis

STATA version 12 and 13 (STATA Corporation, College Station, TX) were used to analyse the data. Statistical significance was tested at the 5% significance level.

2.6.1 Descriptive analysis

The population variables were described and summarised using the appropriate method. Categorical variables were summarised using frequency tables and graphs.

Age was the only continuous variable and it was summarised using its median and range because of it being highly skewed. The number of missing values and the percentage were reported for each variable.

The data for women who returned for the enrolment visit was compared with the data for women who did not return to assess for selection bias was present. Associations between BV and the various characteristics were examined using the Chi-squared test of proportions for categorical variables and Fisher's Exact test for variables with expected values of less than 5 observations for at least 20% of the cells.

2.6.2 Logistic Regression Analysis

Logistic regression was used because the outcome variable of interest was binary. All explanatory variables that had more than 10% of their observations missing were excluded from the logistic analysis. Literature presents various recommendations for handling missing data and most of these methods assume less than 5% of missing values (52, 53). When variables have more than 10% of their values missing the imputation methods may give rise to bias in the result.

All the explanatory variables were first placed in univariate models and regressed on BV. All the variables that had a p-value less than or equal to 0.250 in the univariate analysis were then used to build a multiple variable logistic regression model. A stepwise method was used to build the model. Beginning with the explanatory variable with the least p-value, the variables were added to the model one at a time and the likelihood ratio test used to test if each variable added made the model significantly better.

Changes in the p-values were also noted at each addition. Variables that were not significant at 5% significance level in the final model were dropped out.

Confounding was investigated using marginal effects analysis. The 'chest' command in STATA was used for this analysis which plots and displays the percentage change in estimates for selected exposure variables with the other independent variables as potential confounders.

Interactions were tested for between HIV and other STIs, and between the level of education and employment status. Sensitivity analysis was carried out to examine influential observations and assess the effect of removing these influential observations on the estimates. Post-estimation diagnostic procedures were also carried out to assess model specification and goodness of fit.

2.6.3 Structural Equation Modelling

Structural equation modelling was used to determine the direct and indirect relationships between the underlying and proximate variables with BV infection, and the effects of the variables on each other. The conceptual framework in Figure 2.1 was used to determine which pathways would be assessed.
Demographic and socio-economic factors were considered to be the underlying factors while sexual behavioural, contraceptive use and the presence of STIs were considered to be proximate factors.

A recursive model was specified in two steps using maximum likelihood estimation with robust standard errors. Firstly, all the explanatory variables that were significant at the 5% level in the multiple logistic regressions were tested for their

direct effect on the outcome. From these, all the variables that had a significant direct effect were then used to test for indirect factors related to them as the endogenous variables using the framework in Figure 2.1 to guide specification of the model.

Dummy variables were created for variables with more than two categories and only the statistically significant categories were included in the SEM model. Variables that did not have a significant direct effect on BV were not treated as endogenous variables in the model. The framework in Figure 2.1 guided the specification of the model. Some variables were added a priori to the model based on literature evidence of their association with BV.

The graphical user interface (GUI) in STATA version 13's SEM was used to perform the modelling. All pathways that were not significant at the 5% level were removed from the model.

Equation goodness-of-fit statistics were computed for the model. The coefficient of determination (R^2) statistic was computed to assess model fit. The model was re-specified with default standard errors to assess robustness of the model fit.

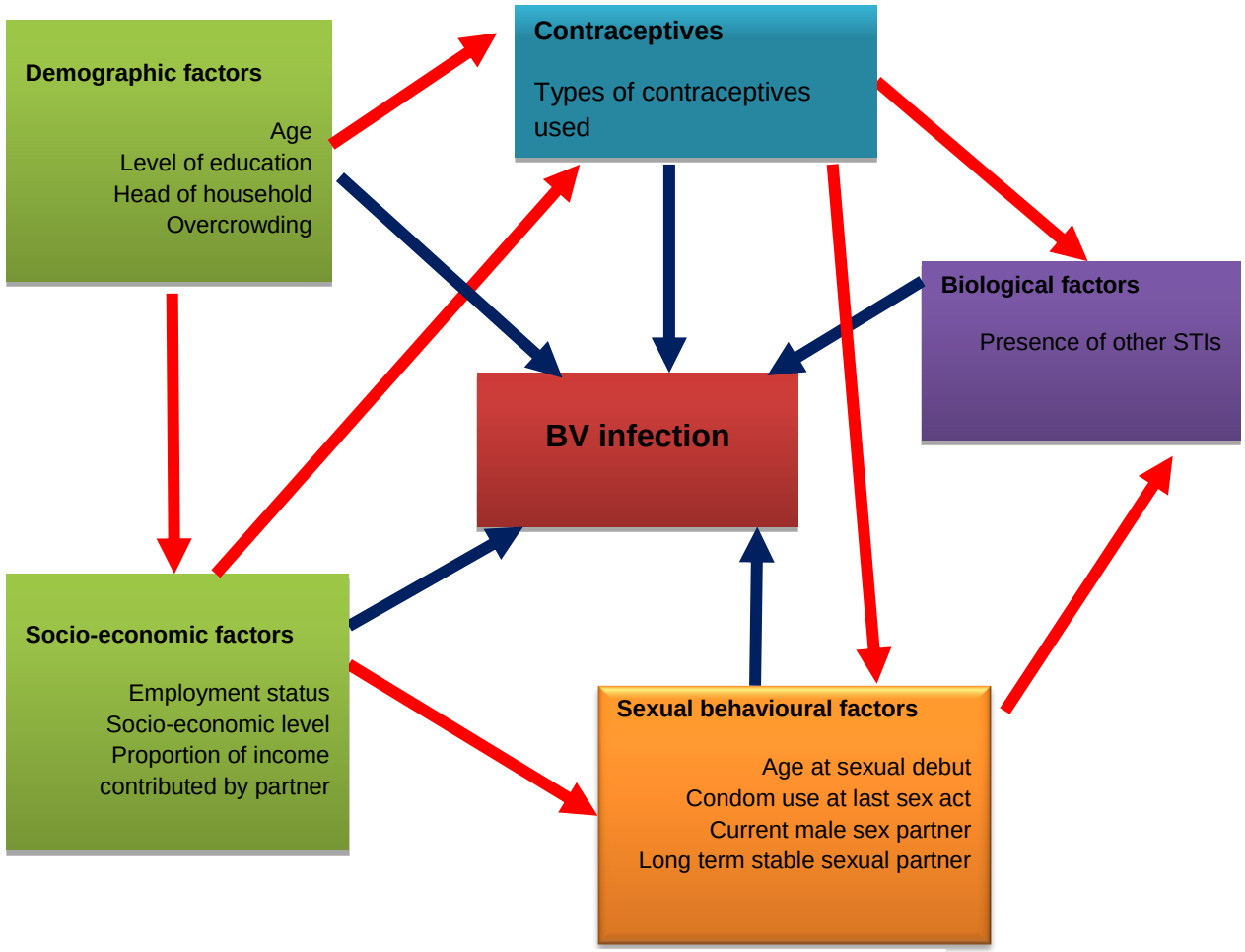


Figure 2.1 Conceptual Framework for Structural Equation Model

Key: Red box – outcome; **Green box-** demographic and socio-economic variables; **Orange box-** sexual behavioural; **Blue box –** contraceptive use variables; **Purple box –** biological variable; **Blue arrow –** direct effect on BV; **Red arrow-**indirect effect on BV

2.7 Ethical Considerations

The primary study was carried out according to international good clinical practice (ICH GCP) guidelines (45). Ethics approval for the primary study was given by the research ethics committee at the University of the Witwatersrand (approval number M050108) (45). The MDP 301 protocol was also reviewed by the National Health Research Ethics Committee of South Africa, and two ethics committees in the United Kingdom. The Food and Drug Administration in America manufactured the gel and also reviewed the protocol. Ethical approval for this study was sought and obtained from the research ethics committee at the University of the Witwatersrand (approval number M120935) and the ethics approval certificate is attached at the end of this report. Permission to use the MDP 301 data was obtained from the WRHI prior to commencement of the study.

CHAPTER 3: RESULTS

Introduction

The results of the study are presented in this chapter. A description of the study sample opens the chapter followed by a summary of the characteristics of the women. Associations between the explanatory variables and BV are presented as well as univariate and multiple variable logistic regression results. The SEM results together with the model are reported.

3.1 Study Participants

A total of 3 825 women were screened between October 2005 and August 2008. The women who returned for the enrolment visit and were enrolled into the study were 2 544 (66.5%). Of those who were enrolled, 50 did not have a BV result and 24 had a BV result graded using a method other than the Hay-Ison method leaving a sample of 2 470 which was used for the analysis. Figure 3.1 is a flow chart of the participants.

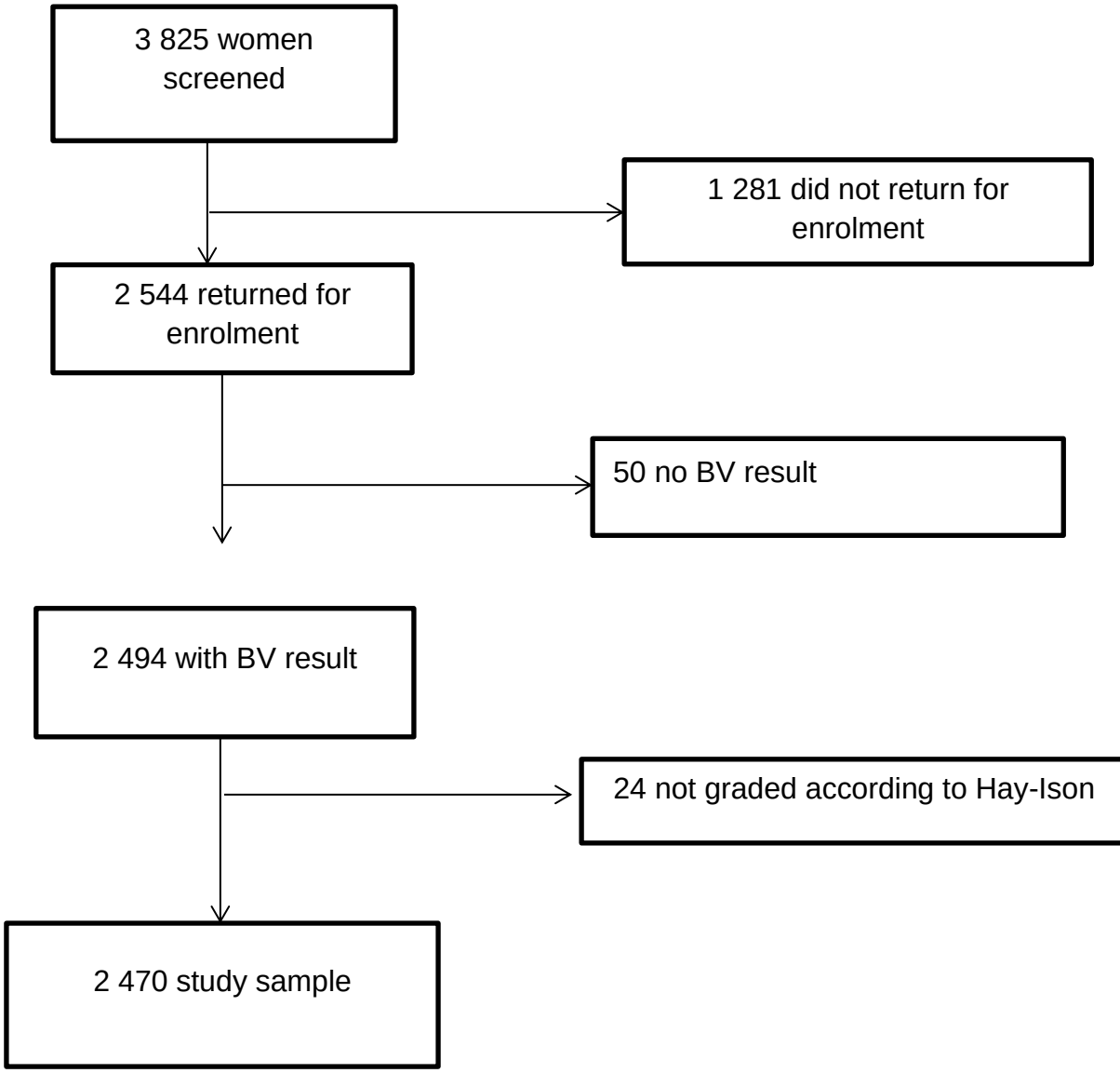


Figure 3.1 Flowchart of study participants

3.2 Characteristics of the women

Table 3.1 summarises the characteristics of the women included in the study and the percentage of missing values for each variable. The missing values were calculated separately from the categories of the variable i.e. using different denominators. The majority of the women were young, with 53.3% of the sample aged between 20 and 30 years and about 90% of them being below the age of 40.

The median age was 25 (IQR 21-33 years). The prevalence of bacterial vaginosis, the outcome of interest, was 40.5%.

Most of the women had at least a secondary school education and 72% reported being unemployed. 51.8% were in the lowest socio-economic level. There was no evidence of overcrowding with over 99% of the women reporting only 1 or two people sleeping per bedroom.

The most commonly used family planning methods were the condom (41.3%) and the injectables (Nur-Isterate – 22.3% and Depo Provera – 12.9%). There were no women using foam, the diaphragm, other injectables, Norplant and oral and vaginal traditional methods therefore they were left out of the analysis.

Condom use at last sex act was high at 61.1%. A large number of the women (86.2%) had their sexual debut below the age of 20 years. About 97% reported being in a stable long-term relationship.

The STI with the highest prevalence was HSV2 (46.4%) followed by HIV at 17.0%. The prevalence of syphilis was 3.1%; 8.2% for *T. vaginalis* infection; 3.1% for *N. gonorrhoeae* infection and 12.7% for chlamydia infection.

Table 3.1 Characteristics of the Women

Characteristic		N	%
Bacterial vaginosis (N= 2 470)			
	No	1 470	59.5
	Yes	1 000	40.5
	Missing	0	0.0
Age			
	<20	284	11.5
	20-30	1,317	53.3
	30-40	606	24.5
	>40	263	10.7
	Missing	0	0.0

Education level (N=2 470)		
Primary and less	139	5.6
Complete/Incomplete secondary	2 177	88.1
Complete/Incomplete Tertiary/vocational	154	6.2
Missing	0	0
Employment status (N=2 469)		
Full time	160	6.5
Part time	193	7.8
Student	215	8.7
Work-seeker	124	5.0
Unemployed	1 777	72.0
Missing	1	0.04
Head of household (N=2470)		
Self	360	14.6
Partner	431	17.5
Sibling	84	3.4
Parent	1 190	48.2
Other relative	405	16.4
Missing	0	0.0
Number of people sleeping per room (N=2 469)		
1-2 people	2 467	99.9
More than 2 people	2	0.01
Missing	1	0.04
Socio-economic level (N=2 470)		
Low	1 280	51.8
Middle	1 110	45.0
High	80	3.2
Missing	0	0.0
Proportion of income contributed by partner (N=2 470)		
Nothing	943	38.2
One third	518	21.0
One half	365	14.8
Three quarters	227	9.2
All	417	16.8
Missing	0	0.0
Family planning -natural rhythm (N= 2 470)		
No		
Yes	2,452	99.3
Missing	18	0.73
	0	0.0
Oral contraceptive pills (N=2 470)		
No	2,200	89.1
Yes	270	11.0
Missing	0	0.0

Diaphragm (N=2 470)			
	No	2,469	99.96
	Yes	1	0.04
	Missing	0	0.0
Injectable, Nuristerate (N=2 470)			
	No	1,919	77.7
	Yes	551	22.3
	Missing	0	0.0
Injectable, Depo Provera (N=2 470)			
	No	2,151	87.1
	Yes	319	12.9
	Missing	0	0.0
Intra-uterine device (N=2 470)			
	No	2,463	99.7
	Yes	7	0.3
	Missing	0	0.0
Condom (N= 2 470)			
	No	1,450	58.7
	Yes	1,020	41.3
	Missing	0	0.0
Traditional, other (N=2 470)			
	No	2,465	99.8
	Yes	5	0.2
	Missing	0	0.0
Other			
	No	2,465	99.8
	Yes	5	0.2
	Missing	0	0.0
Current male sex partner (N=2 470)			
	No	9	0.4
	Yes	2,461	99.6
	Missing	0	0.0
Long term stable partner (N=2 470)			
	No	77	3.1
	Yes	2,393	96.9
	Missing	0	0.0
Condom use at last sex act (N=2 468)			
	No	960	38.9
	Yes	1,508	61.1
	Missing	2	0.08

Age at sexual debut (N= 2469)		86	3.5
	<15 years	2,043	82.7
	15-19 years	340	13.8
	20 years or older	1	0.04
	Missing		
Vaginal insertion (N= 1 652)			
	No	1,644	99.5
	Yes	8	0.5
	Missing	818	33.0
Number of times of vaginal insertion (N=7)			
	More than once a day	3	42.9
	Once a day	3	42.9
	Less than once a day but more than once a week	1	14.3
	Missing	2 463	99.7
Anal sex in past 4 weeks (N=1 652)			
	No	1,624	98.3
	Yes	28	1.7
	Missing	818	33.1
Condom use during anal sex (N= 28)			
	Always	15	53.6
	Most of the time	1	3.5
	Sometimes	5	17.9
	Never	7	25.0
	Missing	2 442	98.9
HIV (N= 2 457)			
	No	2,028	82.5
	Yes	418	17.0
	Discordant	11	0.45
	Missing	13	0.5
Syphilis (N=2 447)			
	No	2,370	96.9
	Yes	77	3.1
	Missing	23	0.9
HSV 2 (N=2 466)			
	No	1,147	46.5
	Yes	1,145	46.4
	Equivocal	174	7.1
	Missing	4	0.2
Trichomonas (N=2 455)			
	No	2,254	91.8
	Yes	201	8.2
	Missing	15	0.6
Neisseria (N=2 460)			
	No	2,358	95.9
	Yes	77	3.1
	Inhibition	25	1.0
	Missing	10	0.4

Chlamydia (N=2 460)			
	No	2,125	86.4
	Yes	313	12.7
	Inhibition	22	0.9
	Missing	10	0.4

Variables with categories with counts of less than 10 that could not be combined with other categories and variables with more than 10% of their values missing were omitted from further analysis. Literature suggests that when a variable has ten or less observations altogether or in some cells the regression coefficients become biased (52, 54). Even though the events per variable (EPV) is not the only factor to be considered for accuracy of parameters, it is still an important consideration (55). The omitted variables were condom use during anal sex, vaginal insertion, number of times of vaginal insertion, family planning-diaphragm, intrauterine device, other family planning methods, other traditional family planning methods and number of people sleeping per room (overcrowding).

3.3 Comparison of Missing and Non-Missing BV Data

Of the women who were screened into the MDP 301 study, 1 281 of them did not return for the enrolment visit. 50 of those who returned did not have a BV result and 24 had a BV result not graded using the Ison and Hay method. Figure 3.2 compares the women with and without BV data on some variables that were significantly associated with BV. The demographic and socio-economic characteristics of the women with missing BV data were compared with those used in the analysis using percentages and the results are presented in Figure 3.3.

The ages of the two groups of women were not different from each other with the median ages being 25 years (range 18-58 years) for both groups. The group with BV data had noticeably more unemployed women than those who did not. This may be because unemployed women had time to return for the enrolment visit during the day. There were more women in the low socio-economic level who did not have BV data.

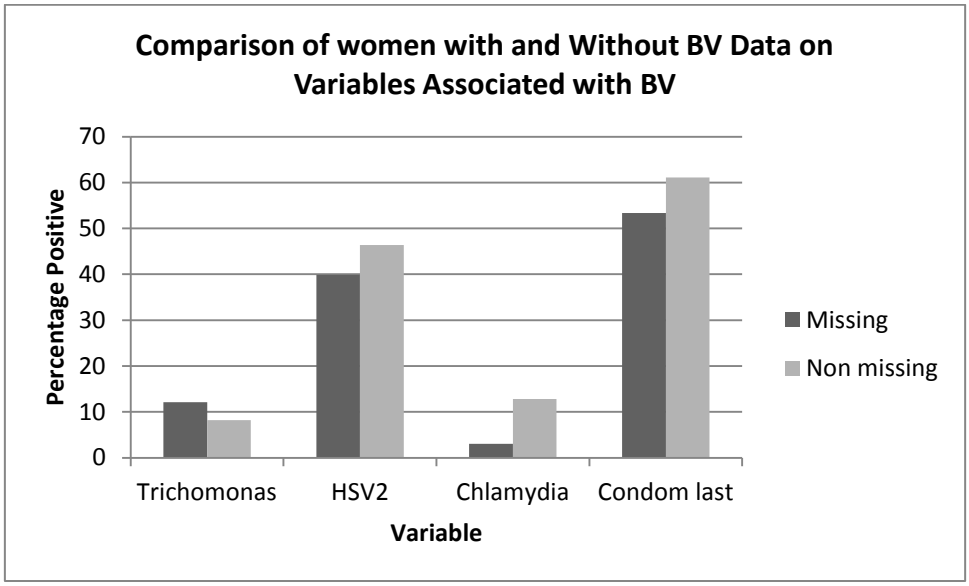


Figure 3.2 Comparison of Women With and Without BV Data on Variables Significantly Associated with BV Infection

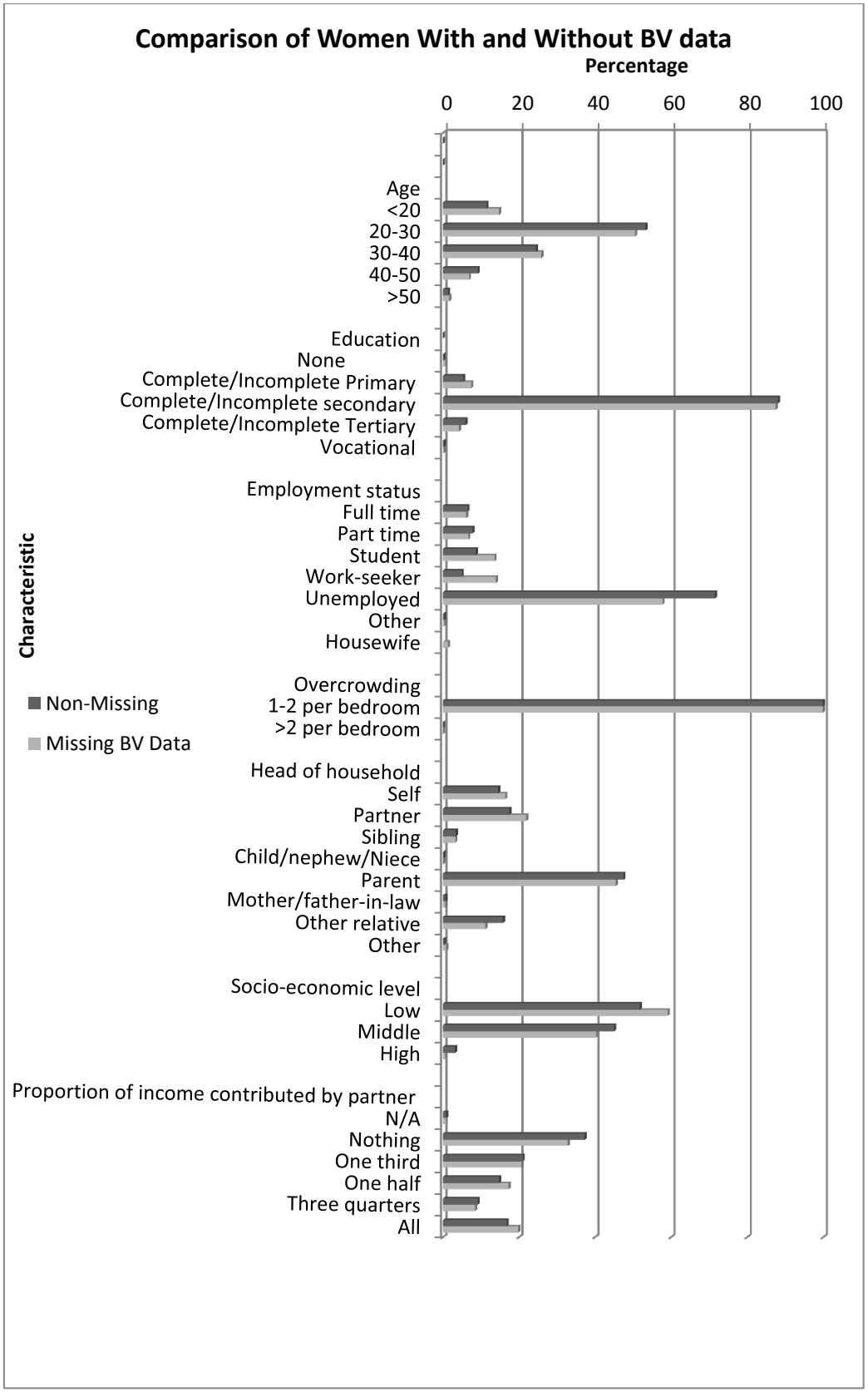


Figure 3.3 Comparison of Women with and Without BV Data

3.4 Bivariate Analysis of Factors Associated With Bacterial Vaginosis

The associations between BV and the characteristics of the women were investigated. Table 3.2 presents the characteristics of the women according to their BV result as well as the results of the significance tests.

Table 3.2 Factors Associated with Bacterial Vaginosis

Characteristic		BV Negative (N=1 470) n (%)	BV Positive (N= 1 000) n (%)	p-value*
Age				
	<20	176 (12.0)	108 (10.8)	
	20-30	759 (51.6)	558 (42.4)	
	30-40	368 (25.0)	238 (23.8)	
	>40	167 (11.3)	96 (9.6)	0.191
Education level				
	Primary or less	80 (5.4)	59 (5.9)	
	Complete/Incomplete secondary	1 294 (88.0)	883 (88.3)	
	Complete/Incomplete tertiary/Vocational	96(6.5)	58(5.8)	0.692
Employment status				
	Full time	96 (6.5)	64 (6.4)	
	Part time	109 (7.4)	84 (8.4)	
	Student	138 (9.4)	77 (7.7)	
	Work-seeker	77 (5.2)	47 ((4.7)	
	Unemployed	1 050 (71.4)	727 (7.3)	0.528
Head of household:				
	Self	228 (15.5)	132 (13.2)	
	Partner	249 (16.9)	182 (18.2)	
	Sibling	44 (3.0)	40 (4.0)	
	Parent	710 (48.3)	480 (48.0)	
	Other relative	239 (16.3)	166 (16.6)	0.337
Proportion of income contributed by partner				
	All	247 (16.8)	170 (17.0)	
	Nothing	564 (38.4)	379 (37.9)	
	One third	307 (20.9)	211 (21.1)	

One half	213 (14.5)	152 (15.2)	
Three quarters	139 (9.5)	88 (8.0)	0.969
Socio-economic level			
Low	757 (51.5)	523 (52.3)	
Middle	674 (45.9)	436 (43.6)	
High	39 (2.6)	41 (4.1)	0.101
Family planning- natural rhythm			
No	1 463(99.5)	989 (98.9)	
Yes	7 (0.5)	11(1.1)	0.074
Oral contraceptive pills			
No	1 322(89.9)	878 (87.8)	
Yes	148 (10.1)	122 (12.2)	0.096
Injectable, Nuristerate			
No	1 154 (78.5)	765 (76.5)	
Yes	316 (21.5)	235 (23.5)	0.240
Injectable, Depo Provera			
No	1 269 (86.3)	882 (88.2)	
Yes	201 (13.7)	118 (11.8)	0.173
Condom			
No	862 (58.6)	588 (58.5)	
Yes	608 (41.4)	412 (41.2)	0.937
Sterilisation			
No	1 415 (96.3)	957 (95.7)	
Yes	55 (3.7)	43 (4.3)	0.485
Current male sex partner:			
No	7 (0.5)	2 (0.2)	
Yes	1 463 (99.5)	998 (99.8)	0.327
Long term stable partner:			
No	49 (3.3)	28 (2.8)	
Yes	1 421 (96.7)	972 (97.2)	0.454
Condom use at last sex act:			
No	543 (36.9)	417 (41.8)	
Yes	927 (63.1)	581 (58.2)	0.015
Age at first sex:			
<15 years	52 (3.5)	34 (3.4)	
15-19 years	1 239 (84.3)	804(80.5)	
20 years and older	179 (12.2)	161 (16.1)	0.021
Anal sex in past 4 weeks			
No	951 (97.9)	673 (98.8)	
Yes	20 (2.1)	8 (1.2)	0.170
HIV			
Negative	1 205 (82.4)	823 (82.7)	
Positive	251 (17.2)	167 (16.8)	
Discordant	6 (0.4)	5 (0.5)	0.919

HSV2		725 (49.5)	422 (42.2)	
	Negative	638 (43.5)	507 (50.7)	
	Positive	103 (7.0)	71 (7.1)	0.001
	Equivocal			
Syphilis RPR				
	Negative	1 411 (97.1)	959 (96.5)	
	Positive	42 (2.9)	35 (3.5)	0.380
Trichomonas				
	Negative	1 389 (95.1)	865 (86.9)	
	Positive	71 (4.9)	130 (13.1)	0.000
Neisseria				
	Negative	1 415 (96.6)	943 (94.7)	
	Positive	35 (2.4)	42 (4.2)	
	Inhibition	14 (1.0)	11 (1.1)	0.035
Chlamydia				
	Negative	1 315 (89.8)	810 (81.3)	
	Positive	136 (9.3)	177 (17.8)	
	Inhibition	13 (0.9)	9 (0.9)	0.000

*****Chi-square test or Fisher's exact test where expected frequencies are less than 5

All the demographic and socio-economic variables were not significantly associated with having BV infection. Using at least one method of family planning was significantly associated with BV infection but none of the separate methods had a significant association. Of the sexual behaviour variables, condom use at last sex act and age at sexual debut were significantly associated with BV.

HIV infection and syphilis had no significant association with BV infection and all the other STIs i.e. *T.vaginalis*, chlamydia, HSV2 and *N. gonorrhoea* were significantly associated with BV.

3.5 Univariate and Multiple-Variable Analysis of Factors Associated with BV

All variables with less than 10% of their values missing were included in the univariate analysis.

The variables that were omitted from this analysis were similar to those omitted in the bivariate analysis. In order to quantify the results from the bivariate analysis, univariate logistic regression included all variables from the bivariate analysis except those indicated above. The bivariate analysis gave an indication of the association of the factors before adjusting for other variables. All variables with a p-value less than or equal to 0.250 in the univariate analysis were then included in the multiple-variable model. Table 3.3 presents the unadjusted (OR) and adjusted odds ratios (AOR) and the corresponding p-values.

Women in the highest socio-economic level showed a trend of being 1.7 times more likely to have BV than those in the lowest level after adjusting for other factors (p=0.033).

Using a condom during the last sex act reduced the chances of having BV infection both in univariate (OR=0.82, 95% CI 0.70-0.96) and multiple-variable analysis (AOR=0.82, 95% CI 0.70- 0.98) and both were highly significant.

Being HIV positive had no significant association with BV infection in this analysis, and neither did a positive syphilis RPR result. Women with *N. gonorrhoeae* infection were 1.8 times more likely to have BV infection (p=0.012) in univariate analysis but there was no significant relationship after adjusting for other factors. Those who had Herpes Simplex Virus (HSV) 2 infection were 1.31 times more likely to have BV after adjusting for other factors (p=0.002).

The adjusted odds ratio for the association of T*. vaginalis* and BV infection was 2.68 (p<0.001). Women with chlamydia infection were 2.05 times more likely to have BV infection compared to those who were not infected after adjusting for other variables (p<0.001).

No significant interactions were found between HIV and other STIs, and between the level of education and employment status.

Table 3.4 Univariate and Multiple-variable Analysis

Variable	Unadjusted OR (95% CI)	p-value	Adjusted OR (95% C I)	p-value
Demographic factors				
Age				
< 20 years	1.00			
20-30 years	1.20 (0.92-1.56)	0.179		
30-40 years	1.05 (0.79-1.41)	0.722		
>40 years	0.94 (0.66-1.32)	0.712		
Education level				
None/Primary	1.00			
Secondary	0.93(0.65-1.31)	0.661		
Tertiary/Vocational	0.82 (0.51-1.31)	0.404		
Head of household				
Self	1.00			
Partner	1.26 (0.95-1.68)	0.112		
Sibling	1.57 (0.97-2.53)	0.065		
Parent	1.17(0.91-1.49)	0.212		
Other relative	1.19 (0.90-1.60)	0.221		
Socio-economic factors				
Employment status				
Full time	1.00			
Part time	1.16 (0.76-1.77)	0.504		
Student	0.84 (0.55-1.28)	0.408		
Work-seeker	0.92 (0.57-1.48)	0.719		
Unemployed	1.04 (0.75-1.44)	0.822		

Socio-economic level				
Low	1.00			
Middle	0.94 (0.80-0.12)	0.432		
High	1.52(0.96-2.38)	0.072	1.66(1.04-2.64)	0.033
Proportion of partner's contribution of income				
All	1.00			
Nothing	0.98 (0.77-1.23)	0.842		
One third	1.00 (0.77-1.30)	0.992		
One half	1.04 (0.78-1.38)	0.804		
Three quarters	0.92 (0.66-1.28)	0.621		
Contraceptive use factors				
Family planning- Natural rhythm				
No	1.00			
Yes	2.32 (0.90-6.02)	0.082		
Oral contraceptive pills				
No	1.00			
Yes	1.24 (0.96-1.60)	0.09		
Family planning- Nur-Isterate				
No	1.00			
Yes	1.21 (0.93-1.36)	0.241		
Family planning-Depo Provera				
No	1.00			
Yes	0.84 (0.66-1.08)	0.173		
Family planning – Condom				
No	1.00			
Yes	0.99 (0.84-1.17)	0.937		
Sexual behavioural factors				
Current male sex partner				
No	1.00			
Yes	2.39 (0.49-11.52)	0.278		

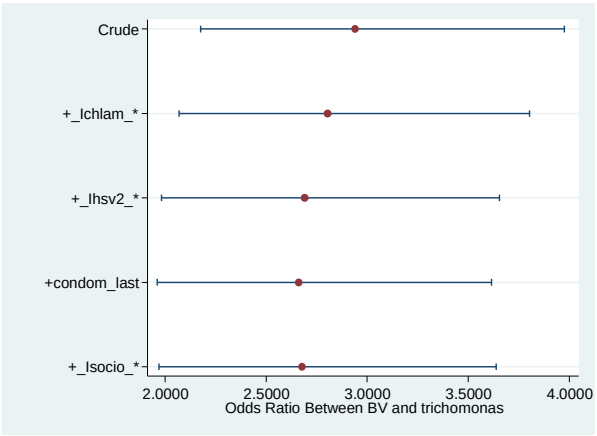
Long term stable partner					
	No	1.00			
	Yes	1.20 (0.75-1.92)	0.455		
Condom use last sex act					
	No	1.00		1.00	
	Yes	0.82 (0.70-0.96)	0.015	0.82 (0.69- 0.97)	0.025
Age at sexual debut					
	<15 years	1.00			
	15-19 years	0.99 (0.64-1.54)	0.973		
	20+ years	1.38 (0.85-2.23)	0.195		
Biological factors					
HIV status					
	Negative	1			
	Positive	0.97 (0.79-1.21)	0.811		
	Equivocal	1.22 (0.37- 4.01)	0.743		
Hsv2					
	Negative	1.00		1.00	
	Positive	1.34(1.13-1.58)	0.001	1.31 (1.10-1.56)	0.002
	Equivocal	1.17 (0.84-1.63)	0.353		
Syphilis RPR					
	Negative	1.00			
	Positive	1.19 (0.75-1.88)	0.469		
Trichomonas					
	Negative	1.00			
	Positive	2.90 (2.14-3.93)	0.000	2.68 (1.97-3.64)	0.000
Neisseria					
	Negative	1.00			
	Positive	1.80 (1.14-2.86)	0.012		
	Inhibition	1.17 (0.53-2.60)	0.692		
Chlamydia					
	Negative	1.00		1.00	
	Positive	2.05(1.62-2.63)	0.000	2.02(1.61-2.62)	0.000
	Inhibition	1.12(0.48-2.62)	0.800		

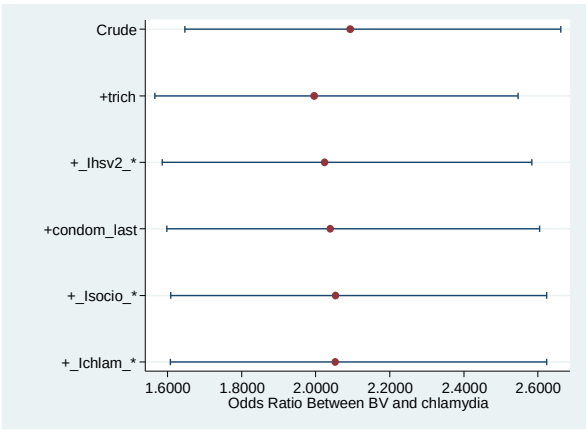
3.5.1 Post-estimation model diagnostics

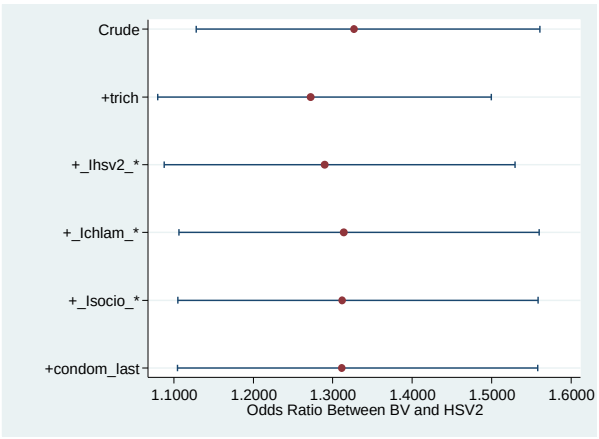
The link test for specification error showed that the regression model was well specified. The Hosmer-Lemeshow test for goodness of fit showed reasonable fit (p=0.5155).

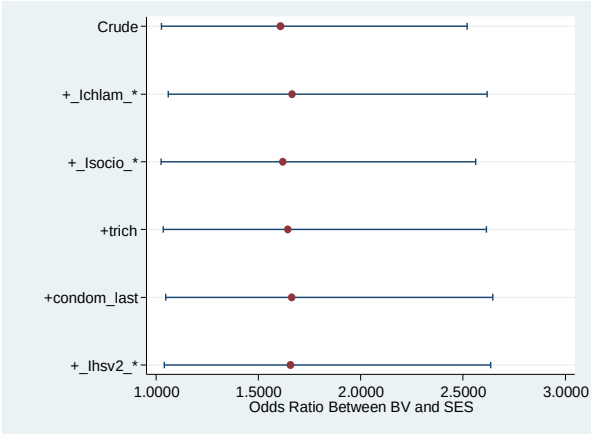
The effect of influential observations was tested using standardised residuals and deviance residuals and no significant change in model estimates were observed after removing the influential observations. No significant collinearity was found between the variables in the model.

No significant confounding observed between the independent variables in the final model. The highest change in estimate was 3.5% change in the odds ratio between high socio-economic level and BV due to the potential confounding effect of being chlamydia positive. Figure 3.4 displays the plots showing the changes in estimates of each independent variable in the final logistic regression model caused by the confounding effects of the other independent variables.


Crude
+_lchlam_*
+_lhsv2_*
+condom_last
+_lsocio_*
2.0000
2.5000
3.0000
3.5000
4.0000
Odds Ratio Between BV and trichomonas


Crude
+trich
+_lhsv2_*
+condom_last
+_lsocio_*
+_lchlam_*
1.6000
1.8000
2.0000
2.2000
2.4000
2.6000
Odds Ratio Between BV and chlamydia


Crude
+trich
+_lhsv2_*
+_lchlam_*
+_lsocio_*
+condom_last
1.1000
1.2000
1.3000
1.4000
1.5000
1.6000
Odds Ratio Between BV and HSV2


Crude
+_lchlam_*
+_lsocio_*
+trich
+condom_last
+_lhsv2_*
1.0000
1.5000
2.0000
2.5000
3.0000
Odds Ratio Between BV and SES

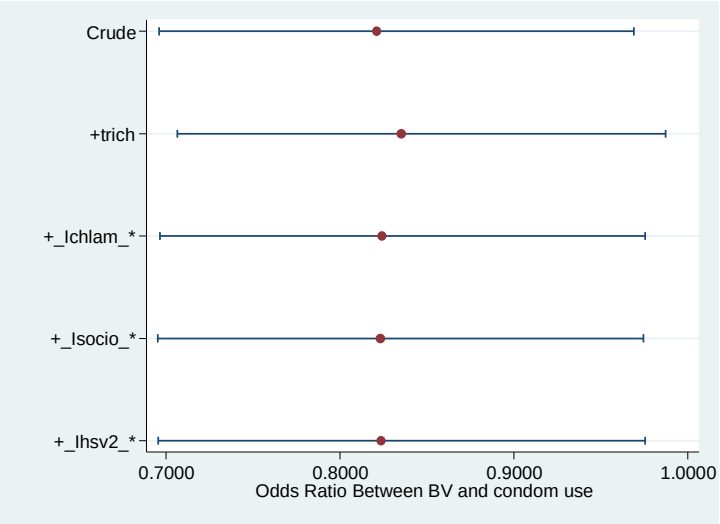


Figure 3.4 Marginal Analysis Plots for Confounding

3.6 Structural Equation Modelling Analysis

The direct, indirect and total effects of the explanatory variables on BV infection were tested, as well as their effects on each other.

Hormonal contraceptive variables were added to the SEM a priori due to the evidence in literature of their effect on BV (23, 30, 31) and on condom use (56, 57). Table 3.4 and Figure 3.5 display the results of the analysis. The connected arrows in the SEM framework in Figure 3.5 show direct and indirect pathways that were statistically significant. A direct pathway is present when an arrow points straight to BV from variable A or B. An indirect pathway is present when an arrow points from variable A to B then to BV the outcome, and B is said to be mediating the effect of A on the outcome .

The effect or relationship is shown by the size and direction of the logit regression coefficient. For indirect effects the beta coefficients from one pathway are multiplied to obtain the coefficient for that pathway.

If more than one indirect pathway exists the coefficients are added together for all the possible pathways to get the total indirect effect. The total effect of a variable on the outcome is obtained by adding together the coefficients for all the direct and indirect pathways. The odds ratios were obtained by calculating the exponent of the beta coefficient. The SEM results displayed and reported are all adjusted for the other endogenous and exogenous variables.

The third column of Table 3.4 shows the variables that had a significant direct association with BV infection. The direct associations, i.e. the variables, are the same as the associations obtained from the logistic regression multiple-variable model. SEM unlike logistic regression is able to extend further and explore indirect associations between the variables simultaneously. The fourth column gives the indirect associations and the last column shows the total effect of each variable.

High socio-economic level was the only underlying factor that had a significant direct effect on BV infection, the effect being positive (beta=0.12, OR=1.1). Condom use at last sexual encounter had a negative direct effect on BV infection, meaning that women who used a condom during their last sexual encounter were less likely to have BV than those who did not use one (beta=-0.043, OR = 0.96).

Out of the three STIs that showed a direct effect on BV, *T. vaginalis* infection had the largest effect (beta=0.24, OR=1.3).

Using Depo Provera, Nur-Isterate and oral contraceptive pills compared to all other methods both all had a positive direct effect on condom use during the last sex act.

Only use hormonal contraceptives i.e. oral contraceptive pills, Depo Provera and Nur-Isterate had significant indirect effects on BV. These effects were mediated through the use of a condom during the last sex act.

Table 3.4 Direct, Indirect and Total Effects of Underlying and Proximate Variables on BV Infection

Effect of	On	Direct	Indirect	Total
Socio-economic factors **High socio-economic level**	BV infection	0.11 ($2.2x10^{-2}$, 0.23)*		0.11 ($2.2x10^{-2}$, 0.23)*
Contraceptive use factors	BV infection		$6.8x10^{-3}$($4.1x10^{-3}$, $1.4x10^{-2}$)*	$6.8x10^{-3}$($4.1x10^{-3}$, $1.4x10^{-2}$)*
Depo Provera Injectable	Condom use	-0.16(-0.22, $-9.6x10^{-2}$)**		-0.16(-0.22, $-9.6x10^{-2}$)**
Nur-Isterate injectable	BV infection		$5.6x10^{-3}$($2.1x10^{-4}$, $1.1x10^{-2}$)*	$5.6x10^{-3}$($2.1x10^{-4}$, $1.1x10^{-2}$)*-
	Condom use	-0.13(-0.18, -0.081)**		0.13(-0.18, -0.081)**
Oral contraceptive pill	BV infection		$9.1x10^{-3}$($6.5x10^{-4}$, $1.8x10^{-2}$)*	$9.1x10^{-3}$($6.5x10^{-4}$, $1.8x10^{-2}$)*
	Condom use	-0.21(-0.28, -0.15)**		-0.21(-0.28, -0.15)**

Sexual behavioural factors **Condom use at last sex act**	BV infection	$-4.3x10^{-2}$ ($-0.82x10^{-2}$,- $3.4x10^{-2}$)*	$-4.3x10^{-2}$ ($-0.82x10^{-2}$,- $3.4x10^{-2}$)*
Biological factors **T. vaginalis infection**	BV infection	0.24(0.17-0.31)**	0.24(0.17-0.31)**
HSV2 infection	BV infection	$6.1x10^{-2}$ ($2.3x10^{-2}$, 0.1)**	$6.1x10^{-2}$ ($2.3x10^{-2}$, 0.1)**
Chlamydia infection	BV infection	0.17(0.11-0.23)**	0.17(0.11-0.23)**

*****p<0.05; **p<0.01; ^{ns} Not significant

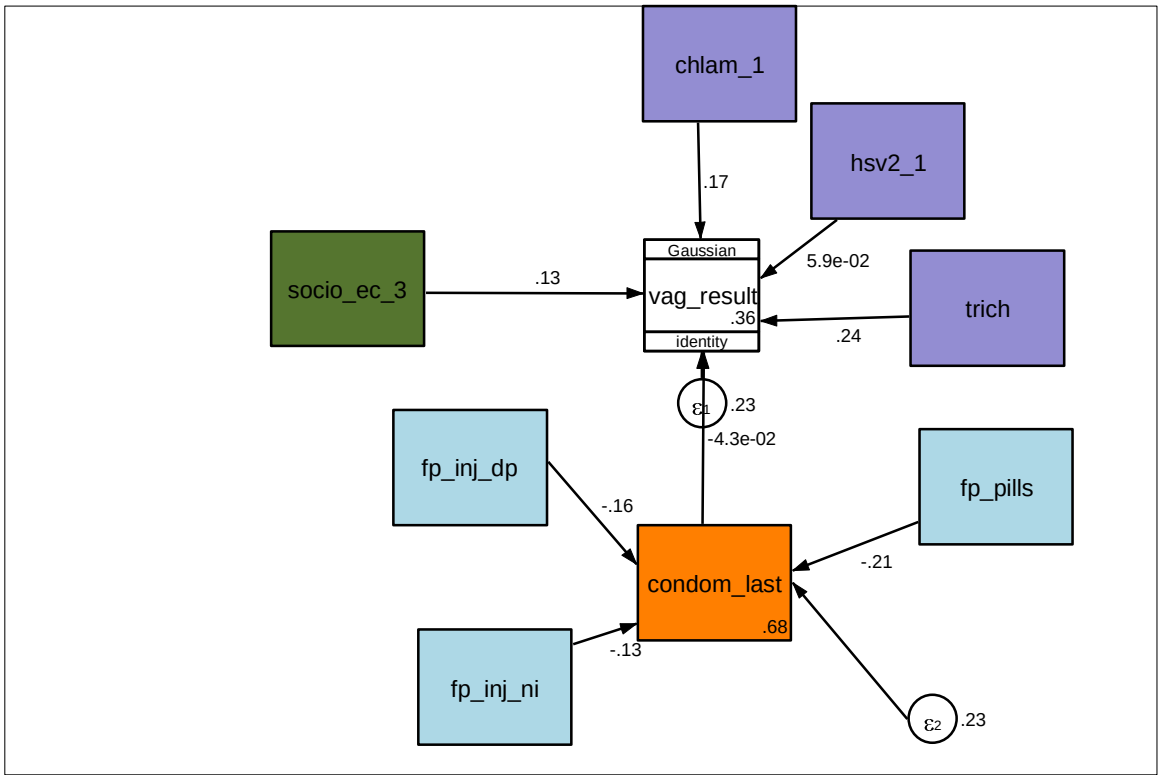


Figure 3.5 Structural Equation Model of the Effects of Underlying and Proximate Variables on BV Infection

Key for variable names: **socio_ec_3 –**high socio-economic level; **fp_inj_dp**-Depo Provera injectable; **fp_inj_ni –** Nur-Isterate injectable; **fp_pills –** oral contraceptive pills; **trich**-trichomonas positive; **chlam_1** – chlamydia positive; **hsv2_1**-HSV2 positive; **vag_result**- BV.

The largest total effect on BV was that of T. vaginalis infection (beta= 0.236, OR=1.3). The largest total effect of any exogenous variable on an endogenous variable other than the outcome was that of using oral contraceptive pills on condom use (beta=-0.154, OR = 0.86).

Condom use at last sex act was the proximate variable mediating the highest number of pathways. The greatest direct effect on BV infection was observed with *T. vaginalis* infection followed by chlamydia infection. These two biological factors also had the largest total effect.

3.6.1 Post-estimation Diagnostics

The equation goodness-of-fit test showed that there might be some significant pathways omitted from the model (R^2=0.042). R^2 indicates the proportion of the variation on the outcome, in this case BV, accounted for by the variables in the model. This statistic however does not take into account the biological relevance of the omitted pathways.

The model was run using normal standard errors and model fit assessed. The model fit test is a likelihood ratio (LR) test of the null hypothesis: the specified model is not different from the population model as measured by the sample. The p-value for the LR test was 0.2434 indicating that the two models are not significantly different and that the specified model reasonably fits the data.

CHAPTER FOUR: DISCUSSION

Introduction

The aim of this study was to explore the underlying and proximate factors associated with bacterial vaginosis and to determine the associations between them with a view to assess the most important pathways. The logistic regression analysis revealed factors that were associated with BV while adjusting for the other factors in the multiple-variable model. The SEM analysis was able to show the factors directly associated with BV as well as enable the postulation of possible indirect relationships through the interaction of the underlying and proximate factors. The results of the analyses are discussed in this chapter together with the limitations of the study and analytical methods used.

4.1 Demographic factors and BV

Both with logistic regression and SEM analysis, none of the demographic variables had an association with BV. An American study was able to demonstrate that BV prevalence varied with age but there was no significant association with the two. These results are similar to what we found (5).

4.2 Socio-economic variables

In univariate analysis, women in the high socio-economic level were 1.5 times more likely to have BV infection compared to those in the low and middle levels. After adjusting for other factors the association was significant. High socio-economic status had a significant direct effect on BV with SEM analysis.

Several studies have also found BV infection to be associated to socio-economic level (5, 6, 9). There were, however no significant indirect pathways from high socio-economic level to BV in the SEM model.

4.3 Contraceptive Use Factors

Previous studies have shown the use of hormonal contraceptives to be protective against BV (30, 56). In this study the use of oral contraceptive pills was marginally significant before adjusting for other factors but other hormonal contraceptives were neither significant in univariate nor multiple variable logistic regression. After adding them a priori to the model these variables (oral contraceptive pills, Nur-Isterate and Depo Provera) had no direct but a positive indirect effect on BV through the mediation of condom use at last sex act. Although the direct effect of condoms on BV indicates that those who use condoms are less likely to have BV, and hormonal contraceptives themselves have been reported to be protective against BV (30), the total indirect effect of hormonal contraceptives on BV through condom use was positive in this study. This may be an indication of interplay between condom use and hormonal contraceptive use that seems to mitigate both their effects on BV. The SEM model further seems to suggest that use of injectable contraceptives negatively impacted on condom use. Injectable contraceptives are longer acting and may be used more by women in long-term relationships, who it has been shown are also less likely to use condoms (58).

Literature has differing views on the association between use of hormonal contraceptives and condoms. The use of hormonal contraceptives has been associated with unprotected sex (i.e. not using condoms) among women (57).

Most of the women in the study reported having a long-term male sexual partner (96.9%) and consistent use of condoms has been inversely associated with long term sexual partners or husbands/wives (58).

4.4 Sexual Behavioural Factors

This study, in agreement with a number of previous studies (2, 9, 18, 59), found in the logistic regression analysis that those who had used condoms during their last sex act were less likely to have BV infection than those who did not (AOR=0.82). The SEM coefficient for the direct effect of condom use on BV showed the same evidence (beta = -0.043).

Condom use as a contraceptive method was however not significantly associated with BV infection in logistic analysis. The difference between the two could be related to consistency of use of condoms during intercourse.

In both logistic and SEM analysis age at sexual debut was not significantly associated with BV infection. Other studies however show that early sexual debut is associated with BV, primarily because age at first intercourse is a predictor of risky sexual behaviour (6). The disparity may be due to incorrect reporting by the women, or recall bias.

4.5 Biological Factors: Presence of STIs

The association between STIs and BV infection has been widely studied and documented. This study found strong univariate and multivariate associations between trichomoniasis, HSV2 and chlamydia infection and BV. *N. gonorrhoeae* infection was only associated with BV infection in the univariate analysis.

The structural equation model also showed significant positive direct effects of these variables on BV. Since this study used observational data it is important to understand the possibility of reverse causality in the interpretation of these results. Results from other studies have shown an association in both directions (1, 6, 9).

The lack of association between BV and HIV infection could be attributed to the large number of women who did not return for the enrolment visit. The primary study also used HIV positivity as an exclusion criterion therefore those who knew their status would not have taken part, regardless of whether they had BV or not.

4.6 Limitations of Study

The strengths and limitations of the study will be discussed under two broad categories namely: measurement of outcome and independent variables, study design and analytical issues.

The data used in this study was collected during the MDP 301 trial using rigorous and sound collection and verification methods. Although about 33% of the women screened did not return for enrolment, it was shown that no significant selection bias was introduced. The laboratory method used to test for BV was a conventional test.

A weakness of secondary data analysis is that investigators are constrained to work with the variables available. The MDP 301 trial used being HIV positive as an exclusion criterion so it is possible that women who already knew their status did not even come for screening.

A cross-sectional design was used for this study and it has the limitation of not being able to show temporal relationships between the outcome and the independent variables.

The conceptual framework used in this study was not used the design stage and collection of data of the primary study. Some useful variables may have been left out of the analysis e.g. number of lifetime sexual partners.

SEM is a fairly novel field in epidemiology and as such there may not be fixed conventions yet on reporting and terminology used. However the technique enabled us in this study to postulate possible pathways between factors associated with BV. SEM is able to adjust further than ordinary logistic regression while finding associations simultaneously. The use of robust standard errors ensured a close model fit.

CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS

5.1 Conclusions

The prevalence of BV in this population of women was 40.5%. None of the underlying variables were found to be associated with BV in logistic analysis but socio-economic level had a significant direct and total effect on BV in the SEM analysis.

Contraceptive use factors were neither significantly associated with BV infection in the multivariable logistic regression nor SEM analysis. They were however indirectly associated with BV through condom use.

Of the sexual behaviour factors, the use of a condom during the last sexual encounter was significantly associated with BV infection with those who used a condom being less likely to have BV.

The presence of *T.vaginalis*, HSV2 or chlamydia infection was significantly associated with BV infection in both analyses.

The mediating effect of condom use and the use of family planning was shown through the use of SEM. Socio-economic status was also shown to have an effect on BV infection through its interaction with other variables like the use of family planning and condom use..

5.2 Scope for further study and recommendations

Lack of education, employment status and poverty are factors that have been shown in other studies to have a significant impact on the risk of women having sexually transmitted or reproductive tract infections. The relationship between BV and high socio-economic status needs further study to understand why women thought to have more knowledge would be more prone to BV.

Reproductive health education and awareness is particularly important in ensuring that the women know how best to combine contraception and prevention of transmission of infection. Educating girls from an early age on sex, sexuality and protecting themselves could help reduce BV infection. The use of condoms needs to continue to be promoted for contraception as well as prevention of infection.

STIs and BV infection have a significant association. Further study needs to be carried out to ascertain the temporal relationship between the two and the factors that increase the risk of getting BV given having an STI or vice versa. The impact of these two on HIV transmission warrants a concerted effort in managing both effectively. It remains unknown also if the organisms that lead to BV depend on the presence of a particular STI.

The results of this study require a follow up with a prospective study to shed more light on the interplay between underlying and proximate factors.

References

1. Baisley K, Changalucha J, Weiss HA, Mugeye K, Everett D, Hambleton I, et al. Bacterial vaginosis in female facility workers in north-western Tanzania: prevalence and risk factors. Sexually Transmitted Infections. 2009;85(5):370.
2. Nwadioha S, Egah D, Banwat E, Egesie J, Onwuezobe I. Prevalence of bacterial vaginosis and its risk factors in HIV/AIDS patients with abnormal vaginal discharge. Asian Pacific Journal of Tropical Medicine. 2011;4(2):156-8.
3. Bradshaw CS, Morton AN, Garland SM, Morris MB, Moss LM, Fairley CK. Higher-Risk Behavioral Practices Associated With Bacterial Vaginosis Compared With Vaginal Candidiasis. Obstetrics & Gynecology. 2005;106(1):105-14
4. Sobel JD. Bacterial vaginosis. Annual Review of Medicine. 2000;51(1):349-56.
5. Allsworth JE, Peipert JF. Prevalence of Bacterial Vaginosis: 2001–2004 National Health and Nutrition Examination Survey Data. Obstetrics & Gynecology. 2007;109(1):114-20
6. Bukusi EA, Cohen CR, Meier AS, Waiyaki PG, Nguti R, Njeri JN, et al. Bacterial vaginosis: risk factors among Kenyan women and their male partners. Sexually Transmitted Diseases. 2006;33(6):361.
7. Sewankambo N, Gray RH, Wawer MJ, Paxton L, McNairn D, Wabwire-Mangen F, et al. HIV-1 infection associated with abnormal vaginal flora morphology and bacterial vaginosis. The Lancet. 1997;350(9077):546-50.
8. Cohen CR, Lingappa JR, Baeten JM, Ngayo MO, Spiegel CA, Hong T, et al. Bacterial Vaginosis Associated with Increased Risk of Female-to-Male HIV-1 Transmission: A Prospective Cohort Analysis among African Couples. PLoS Med. 2012;9(6).
9. Cherpes TL, Hillier SL, Meyn LA, Busch JL, Krohn MA. A delicate balance: risk factors for acquisition of bacterial vaginosis include sexual activity, absence of hydrogen peroxide-producing lactobacilli, black race, and positive herpes simplex virus type 2 serology. Sexually Transmitted Diseases. 2008;35(1):78.
10. Wilgert JHHMvd, Morrison CS, Cornelisse PGA, Munjoma M, Moncada J, Awio P, et al. Bacterial vaginosis and vaginal yeast, but not vaginal cleansing, increase HIV-1 acquisition in African women. Journal of Acquired Immune Deficiency Syndromes. 2008;48(2):203-10.
11. Myer L, Denny L, Telerent R, Souza Md, Wright(Jnr) TC, Kuhn L. Bacterial vaginosis and susceptibility to HIV infection in South African women: a nested case-control study. The Journal of Infectious Diseases. 2005;192:1372-80.
12. Denslow SA, Westreich DJ, Firnhaber C, Michelow P, Williams S, Smith JS. Bacterial vaginosis as a risk factor for high-grade cervical lesions and cancer in HIV-seropositive women. International Journal of Obstetrics and Gynecology. 2011;114:273-77.
13. Cohen CR, Lingappa JR, Baeten JM, Ngayo MO, Spiegel CA, Hong T, et al. Bacterial vaginosis associated with increased risk of female-to-male HIV- 1 transmission: a prospective cohort analysis among African couples. PLoS Medicine. 2012;9(6).
14. Myer L, Denny L, Telerant R, de Souza M, Wright TC, Kuhn L. Bacterial vaginosis and susceptibility to HIV infection in South African women: a nested case-control study. Journal of Infectious Diseases. 2005;192(8):1372-80.
15. Allsworth JE, Peipert JF. Severity of bacterial vaginosis and the risk of sexually transmitted infection. American journal of obstetrics and gynecology. 2011;205(2):113. e1-. e6.

16. Peipert JF, Montagno AB, Cooper AS, Sung CJ. Bacterial vaginosis as a risk factor for upper genital tract infection. American journal of obstetrics and gynecology. 1997;177(5):1184-7.
17. Rugpao S, Sriplienchan S, Rungruengthanakit K, Lamlertkittikul S, Pinjareon S, Werawatakul Y, et al. Risk factors for bacterial vaginosis incidence in young adult Thai women. Sexually Transmitted Diseases. 2008;35(7):643.
18. Fethers Kathrine A., Fairley Christopher K., Hocking Jane S. GLC, S. BC. Sexual risk factors and bacterial vaginosis: a systematic review and meta-analysis. Clinical Infectious Diseases. 208;47:1426-35.
19. Lewis JJC, Donnelly CA, Mare P, Mupambireyi Z, Garnett GP, Gregson S. Evaluating the proximate determinants framework for HIV infection in rural Zimbabwe. Sexually Transmitted Infections. 2007;83(suppl 1):i61-i9.
20. Boerma J, Weir SS. Integrating demographic and epidemiological approaches to research on HIV/AIDS: the proximate-determinants framework. Journal of Infectious Diseases. 2005;191(Supplement 1):S61.
21. Victora CG, Huttly SR, Fuchs SC, Olinto M. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. International Journal of Epidemiology. 1997;26(1):224-7.
22. Pybus V, Onderdonk AB. Microbial interactions in the vaginal ecosystem, with emphasis on the pathogenesis of bacterial vaginosis. Microbes and Infection. 1999;1(4):285-92.
23. Holzman C, Leventhal JM, Qiu H, Jones NM, Wang J. Factors linked to bacterial vaginosis in nonpregnant women. Journal Information. 2001;91(10).
24. Hay P. Bacterial vaginosis. Medicine. 2005;33(10):58-61.
25. Ison C, Hay P. Validation of a simplified grading of Gram stained vaginal smears for use in genitourinary medicine clinics. Sexually Transmitted Infections. 2002;78(6):413-5.
26. Gray RH, Wawer MJ, Brookmeyer R, Sewankambo N, Serwadda D, Wabwire-Mangen F, et al. Probability of HIV transmission per coital act in monogamous, heterosexual discordant couples in Rakai, Uganda. Lancet. 2001;357:1149-53.
27. Morris M, Nicoll A, Simms I, Wilson J, Catchpole M. Bacterial vaginosis: a public health review. British Journal of Obstetrics and Gynaecology. 2001;108:439-50.
28. Morris M, Nicoll A, Simms I, Wilson J, M C. Bacterial vaginosis: a public health review. British Journal of Obstetrics and Gynaecology. 2001;108:439-50.
29. Nilsson U, Hellberg D, Shoubnikova M, Nilsson S, MÅRDH PERA. Sexual behavior risk factors associated with bacterial vaginosis and Chlamydia trachomatis infection. Sexually Transmitted Diseases. 1997;24(5):241.
30. Shoubnikova M, Hellberg D, Nilsson S, Mårdh PA. Contraceptive use in women with bacterial vaginosis. Contraception. 1997;55(6):355-8.
31. Baeten JM, Nyange PM, Richardson BA, Lavreys L, Chohan B, Martin HL, et al. Hormonal contraception and risk of sexually transmitted disease acquisition: Results from a prospective study. American journal of obstetrics and gynecology. 2001;185(2):380-5.
32. Ness RB, Hillier SL, Richter HE, Soper DE, McGregor J, Bass DC, et al. Doucning in relation to bacterial vaginosis, lactobacilli and facultative bacteria in the vagina. Obstetrics and Gynaecology. 2002;100(4):765-72.
33. Gallo M F, Macaluso M, Warner L, Fleenor M E, Hook E W 3rd, Brill I, et al. Bacterial vaginosis, gonorrhea, and chlamydial infection among women attending

a sexually transmitted disease clinic: a longitudinal analysis of possible causal links. Annals of Epidemiology. 2012;22(3):213-20.
34. Yoshimura K, Yoshimura M, Kobayashi T, Kubo T, Hachisuga T, M K. Can bacterial vaginosis help to find sexually transmitted diseases, especially chlamydial cervicitis? Internation Journal of STD and AIDS. 2009;20:108-11.
35. Mosley WH, Chen LC. An analytical framework for the study of child survival in developing countries. Population and Development Review. 1984;10:25-45.
36. Lopman B, Nyamukapa C, Mushati P, Mupambireyi Z, Mason P, Garnett GP, et al. HIV incidence in 3 years of follow-up of a Zimbabwe cohort—1998–2000 to 2001–03: contributions of proximate and underlying determinants to transmission. International Journal of Epidemiology. 2008;37(1):88-105.
37. Amorim LDAF, Fiaccone RL, Santos CAST, Santos TN, Moraes LTLP, Oliveira NF, et al. Structural equation modeling in epidemiology. Cadernos de Saúde Pública. 2010;26(12):2251-62.
38. Sanchez BN, Budtz-Jørgensen E, Ryan LM, Hu H. Structural equation models. Journal of the American Statistical Association. 2005;100(472):1443-55.
39. Beran T, Violato C. Structural equation modelling in medical research: a primer. BioMedCentral Research Notes. 2010;3(267).
40. C S. General intelligence, objectively determined and measured.

. American Journal of Psychology

1904;15:201-93.
41. Wright S. On the nature of size factors. Genetics. 1918;3:367-74.
42. Violato C, Hecker KG. How to use structural equation modeling in medical education research: A brief guide. Teaching and learning in medicine. 2007;19(4):362-71.
43. LP S. STATA 12 Reference Manual. Structural Equation Modelling

12 ed. Texas: Stata Press; 2011. p. Statistical Package.
44. McCormack S, Ramjee G, Kamali A, Rees H, Crook AM, Gafos M, et al. PRO2000 vaginal gel for prevention of HIV-1 infection (Microbicides Development Programme 301): a phase 3, randomised, double-blind, parallel-group trial. The Lancet. 2010;376(9749):1329-37.
45. Nunn A, McCormack S, Crook AM, Pool R, Rutterford C, Hayes R. Microbicides Development Programme: design of a phase III trial to measure the efficacy of the vaginal microbicide PRO 2000/5 for HIV prevention. Trials. 2009;10(1):99.
46. NCSS. Power Analysis and Sample Size. Kaysville Utah, USA: NCSS; 2011.
47. Hsieh F Y BDA, Larsen M D. A simple method of sample size calculation for linear and logistic regression. Statistics in Medicine. 1998;17:1623-34.
48. Hoe SL. Issues and procedures in adopting structural equation modelling technique. Journal of Applied Quantitative Methods. 2008;3(1):76-83.
49. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: a review. The Journal of Educational Research. 2006;99(6):323-38.
50. Garver MS, Mentzer JT. Logistics research methods: employing structural equation modeling to test for construct validity. Journal of Business Logistics. 1988;20(1):33-57.

51. Hay PE, Lamont RF, Taylor-Robinson D, Morgan DJ, Ison C, Pearson J. Abnormal bacterial colonisation of the genital tract and subsequent preterm delivery and late miscarriage. BMJ. 1994;308(6924):295-8.
52. Lynch S M. Missing data (Soc 504). 2003.
53. Johnson D. R., Young R. Towards best practices in analyzing datasets with missing data: comparisons and recommendations. Journal of Marriage and Family. 2011;73:926-45.
54. Peduzzi P., Concato J., Kemper E., Holford T R., Feinstein A R. A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology. 1996;49(12):1373-9.
55. Courvoisier D S., Combescure C., Agoritsas T., A G-A, Perneger T V. Performance of logistic regression modelling: beyond the number of events per variable, the role of data structure. Journal of Clinical Epidemiology. 2011;64:993-1000.
56. Smart S, Singal A, A M. Social and sexual risk factors for bacterial vaginosis. Sexually Transmitted Infections. 2004;80:58-62.
57. Heffron R, Were E, Celum C, Mugo N, Baeten JM. A prospective study of contraceptive use among Africa women in HIV-1 serodiscordant partnerships. Sexually Transmitted Diseases. 2010;37(10):621-8.
58. Agha S. Factors associated with HIV testing and condom use in Mozambique: implications for programs Reproductive Health. 2012;9(1):20-8.
59. Schwebke J. R. , Desmond R. Risk factors for bacterial vaginosis in women at high risk of sexually transmitted diseases. Sexually Transmitted Diseases. 2005;32(11):654-8.

APPENDIX 1: GRAM STAIN METHOD (25)

1. Flood smears with crystal violet for 30 seconds and wash off with tap water.
2. Flood smears with Lugol's iodine for 30 seconds and wash off with tap water.
3. Decolourise the smears for 5-10 seconds with acetone.
4. Counter-stain with 15% basic Fuchsin for 1 minute and rinse off with tap water.
5. Blot dry the smears and examine under oil immersion at x1000 magnification.

ETHICS CLEARANCE

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Ms Mercy Manyema

CLEARANCE CERTIFICATE	**M120935**
PROJECT	Factors Associated with Bacterial Vaginosis in Sexually Active women Enrolled in the Microbial Development Programme (MDP) 301 Trail
INVESTIGATORS	Ms Mercy Manyema.
DEPARTMENT	School of Public Health
DATE CONSIDERED	28/09/2012
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 28/09/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Dr Harry Moultrie

DECLARATION OF INVESTIGATOR(S)
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...