A Risk-Score for Predicting the Presence of Treatable Sexually Transmitted Infections in Kenyan Women Planning Conception

Anne Naipanoi Pulei

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Public Health

University of Washington

2020

Committee:
Raymond Scott McClelland, Chair
John Kinuthia
Erica Lokken
Barbra Richardson

Program Authorized to offer Degree:

Department of Global Health

© Copyright 2020

Anne Naipanoi Pulei

University of Washington

Abstract

A Risk-Score for Predicting the Presence of Treatable Sexually Transmitted Infections in Kenyan Women Planning Conception

Anne Naipanoi Pulei

Chair of the Supervisory Committee:

Raymond Scott McClelland
Departments of Global Health, Medicine and Epidemiology

Treatable sexually transmitted infections (STI) disproportionately affect women's reproductive health and contribute to poor neonatal outcomes. Because these infections are frequently asymptomatic, identifying women at higher risk of having STIs using risk scores may provide a cost-effective screening approach in regions where universal screening of pregnancy planners is not performed. The aim of this study was to determine the prevalence and correlates of *Chlamydia trachomatis, Neisseria gonorrhoeae,* and *Trichomonas vaginalis* among Kenyan pregnancy planners, and to develop a risk score to identify those with a higher likelihood of current chlamydial infection. Kenyan women with fertility intent enrolled into a cohort study between April 2017 and March 2020 contributed data for this cross-sectional analysis. Logistic regression was used to estimate odds ratios of the association between demographic, behavioral, and clinical risk factors and prevalent STIs, both as a group, and for *C. trachomatis* alone. Based on the regression coefficients, prediction models were developed to identify women with increased likelihood of current *C. trachomatis* infection. The most common STI was *C. trachomatis* (51/691, 7.4%); *N.*

gonorrhoeae (5/691, 0.5%), and *T. vaginalis* infections (6/687, 0.9%) were rare. The prevalence of any one or more of these STI was 60/688 (8.7%). Risk factors for any STI included age less than 25 (OR 2.41; 95%CI: 0.69-8.48), partner's age less than 25 (OR 17.22; 95%CI: 3.74-79.24), Alcohol Use Disorders Identification Test (AUDIT) score ≥8 (OR 5.54; 95%CI: 1.61-19.04), and bacterial vaginosis (BV) (OR 2.49; 95%CI: 1.46-4.24). A risk score for predicting *C.trachomatis* infection, ranging from 0-6, derived from the participant's age, AUDIT score, and BV status yielded an area under receiver operating curve (AUROC) of 0.78 (95%CI: 0.72-0.84). Using a score cutoff of 0 versus ≥1, 478/691 (69.2%) were classified as higher-risk for *C. trachomatis* (sensitivity=98.0%, 95%CI: 89.6-100.0; specificity=33.1%, 95%CI: 29.5-36.9). At a higher cutoff of ≤2 versus ≥3, the risk score identified 31.8% of women as higher risk (sensitivity=70.6%, 95%CI: 56.2-71.3, specificity=71.3%, 95%CI: 67.7-74.5). Among women classified as higher risk by the risk scoring tool, the numbers needed to screen using nucleic acid amplification based tests (NAAT) testing to detect one *C. trachomatis* infection were 10.0 for the 0 versus ≥1 cutoff and 7.7 for the higher cutoff of ≤2 versus ≥3. This risk-scoring tool may be useful for identifying higher risk women for *C. trachomatis* screening among Kenyan pregnancy planners. Employing the tool could provide a cost-conscious approach for initiating species-specific testing for *C. trachomatis* infection.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my committee chair, Professor R. Scott McClelland who has journeyed with me from my application into the MPH program, conception of this research and all through my academic and research journey in this period. His mentorship and dedication has been invaluable. I am grateful to my committee members: Dr. John Kinuthia for his guidance & encouragement, Dr. Erica Lokken for her guidance, dedication, and concern for my general well-being, and Professor Barbra Richardson for her guidance, reassurance and assistance in data analysis and the risk score derivation. I hope to be able to provide such dedicated mentorship in my future academic career.

I am grateful to Jillian Pintye and Sara Khor for their help in prediction models and risk scores development.

Special thanks goes to Vice Dean Carey Farquhar, the Director of the International AIDS Research and Training Program (IARTP), Sonora, Gisella, Michelle, and Adrianne for helping us settle in Seattle and for ensuring that our academic processes ran smoothly. I appreciate Sarah Masyuko for her insight, support and encouragement.

I am very grateful to my Global Health cohort (2020) for our time together, encouragement, contagious working spirit and cheerfulness. I also appreciate the faculty and staff at the Department of Global Health for the support and encouragement. I am specifically grateful to Steve, Ahoua, Deepa, Andreas, Noura and Julie for their great involvement in our learning milestones and for their immense support.

I truly appreciate the support I got from my family and friends: I will be forever indebted to the Wiggins (Susan, Woody, Sara, Ellwood, Felice, and Thalia) for their encouragement, love, kindness, and for making this year memorable for me. I appreciate KeliAnne Hara-Hubbard for the skills she shared with me during our practicum that has made my writing for this thesis much easier. I am grateful to my friend and boss Pamela Mandela for her support this year. I am forever grateful to my mother, Mrs. Lois Pulei for her words of encouragement, reassurance and constant love. I am also grateful to my siblings Alex Lengete, Nancy Nampa and Steven Sampao for their encouragement. I am forever indebted to my nephews Kasaine and Lemayian Metto for the joy and renewed energy they rekindled in me whenever I felt fatigued. I truly appreciate the support and love I got from Mr. John Omenge and Mrs. Melissa Morangi, Kokulamuka, Morembe, and Nyabwari Morangi. Special appreciation to Kwegena Simba Morangi, my partner, for his enduring love and invaluable support this year.

I also wish to thank the MPTB study staff and study participants without whom this project would not have been possible.

ASANTENI SANA!

DEDICATION

To my father Francis Pulei Ole Munkush, you will always be missed. *Ashe Oleng'*

To my mother Priscillah Lois Pulei, and my mentors.

INTRODUCTION

Sexually transmitted infections (STIs) are a major public health concern globally.[1] The World Health Organization (WHO) estimates that one million curable STIs are acquired daily.[1,2] The most common treatable STIs are chlamydia (*Chlamydia trachomatis*), gonorrhea (*Neisseria gonorrhoeae*), trichomoniasis (*Trichomonas vaginalis*), and syphilis (*Treponema pallidum*) which account for 376 million new infections annually.[1] These STIs are largely asymptomatic and are often not detected or treated, resulting in reproductive complications.[3] The greatest burden of complications of STIs falls on women, including pelvic inflammatory disease (PID), chronic pelvic pain, ectopic pregnancy, and infertility.[4] Genital infections are also associated with adverse pregnancy outcomes such as stillbirth, preterm labor, congenital and neonatal infections, and low birth weight.[5] Lastly, HIV-seronegative individuals with an STI are at a higher risk of HIV acquisition, while those who are seropositive and have an STI have higher infectivity.[6–9]

Most high-income countries recommend routine screening for STIs as part of preconception care and at the first antenatal clinic (ANC) visit.[5,10] These guidelines also recommend annual STI screening for sexually active women younger than 25 years, high-risk women such as sex workers, persons who inject drugs, and those in correctional facilities.[10,11] Unfortunately, STI screening with sensitive, species-specific nucleic acid amplification based tests (NAAT) is costly and not routinely conducted in low-and middle-income countries (LMICs).[12] Since diagnostic capacity for identifying STIs is limited in these countries, the WHO proposed a syndromic approach to the management of STIs in 2003.[13] This approach has been adopted in in many LMICs, including Kenya.[14,15] Syndromic management involves the identification of a consistent collection of clinical symptoms and easily recognized signs. Treatment then targets the majority of the organisms responsible for producing those clinical features. While this approach does provide a tool kit for treatment of symptomatic patients, syndromic management of STIs does not capture asymptomatic cases or those who may be symptomatic but fail to seek care.

Further, the diagnostic accuracy of syndromic management for chlamydia and gonorrhea is limited, resulting in overtreatment of those who are symptomatic.[12]

Screening pregnant women and women planning pregnancies for treatable STIs in LMICs, while desirable, is considered too costly to implement in the context of competing public health priorities.[12] A risk score may provide a cost-effective method of identifying individuals most likely to benefit from screening for a prevalent treatable STIs and could be implemented in ANCs or in the context of preconception care. While numerous studies have reported on the prevalence and risk factors for STIs in women in sub-Saharan Africa, most focus on high-risk groups including adolescents, women who engage in transactional sex, and women in cohorts studying HIV acquisition and transmission.[16,17] Some factors associated with an increased risk of STIs among women include being young, having multiple sexual partners, previous history of STI, condomless sex, low education status, alcohol use, and injection drug use.[18–21] Little is known regarding treatable STIs in general-population African women who are currently planning to conceive. The objective of this study was to assess the prevalence and correlates of treatable STIs among Kenyan women who are planning to become pregnant, and to develop a risk assessment tool for predicting *C. trachomatis* infection in this population.

METHODS

Study design, setting and population

This cross-sectional study utilized enrollment data from the Microbiota and Preterm Birth (MPTB) study, an ongoing study that is examining the association between preconception vaginal microbiota and the risk of spontaneous preterm birth in Nairobi and Mombasa, Kenya. Detailed methods for the study have been published.[22] Participants included in this analysis were enrolled between April 18, 2017 and March 18, 2020.

The MPTB study enrolled HIV-negative women who were recruited from facilities providing reproductive health care. These include women attending family planning clinics for contraceptive

implant and intrauterine device (IUD) removal to become pregnant. Inclusion criteria for the MPTB study included age ≤45 years, planning to conceive, HIV-negative, planning to remain in the study area through pregnancy, able to provide informed consent, and willing to undergo study procedures as outlined in the consent. For women in discordant couples, the partner was required to have an undetectable HIV viral load or the participant was required to be on pre-exposure prophylaxis (PrEP) to prevent HIV infection. Women were excluded if they were currently pregnant, using contraception other than condoms for HIV/STI prevention, reported having sought care for infertility in the past, or reported a medical history suggestive of having a risk of preterm birth due to known cause.[22] For this analysis, we also excluded women who were missing STI results from enrollment.

The Kenyatta National Hospital/University of Nairobi Ethics & Research Committee and University of Washington's Institutional Review Board approved this study. All participants provided written informed consent.

Study procedures and laboratory methods

After screening for HIV status by a rapid HIV test, pregnancy by a urine pregnancy test, and an eligibility interview, eligible participants were able to enroll. At enrollment, participants underwent a structured interview to capture demographic information, sexual behavior, and reproductive, contraceptive, and medical history. For this analysis, since no participants were on contraception at enrollment, recent contraceptive use was defined as use of any modern, non-barrier contraceptive in the past month. A study clinician performed a physical examination and speculum-assisted pelvic examination that included collection of specimens for STI diagnosis. Participants received treatment for STIs by syndromic management at the enrollment visit. Species-specific STI treatment was provided at the next visit if women had a positive NAAT test and had not received effective treatment based on syndromic management at enrollment.

Saline and potassium hydroxide wet mounts were examined for the presence of motile trichomonads, clue cells, and yeast. Gram stained slides were evaluated for bacterial vaginosis (BV) using the criteria of Nugent and Hillier.[23] Testing for *N. gonorrhoeae, C. trachomatis,* and *T. vaginalis* was performed using NAATs (Aptima, Hologic Corporation, San Diego, CA) according to the manufacturer's specifications.

Statistical analysis

Baseline demographic, behavioral, and reproductive characteristics were summarized using proportions and medians, as appropriate. To assess correlates of any prevalent STI, including *C. trachomatis, N. gonorrhoeae,* and *T. vaginalis*, univariate and multivariate logistic regression were performed to estimate odds ratios and 95% confidence intervals (CI). In addition, regression models were used to identify correlates of *C. trachomatis* infection alone. Variables that yielded a p-value <0.10 in univariate analyses were selected to be included in the multivariate logistic regression model. We assessed collinearity between potential correlates using bivariate correlation and standard errors for logistic regression coefficients.[24] Where collinearity was detected, one variable was selected based on clinical relevance.

Risk score development

A risk score was developed to identify women at higher risk for prevalent *C. trachomatis* due to the higher prevalence in this study population compared to the other STI. Variables included in the final multivariate logistic regression model were included for prediction. For categorical variables, the lowest risk group was set as the referent group to make the risk score easier to implement in a clinical setting. For the continuous age variable, several cutoffs were considered including: <25 years or ≥25 representing the known high-risk age group for *C. trachomatis as* reported in literature ,[21] <27 years or ≥27 selected as the optimal cutoff by area under receiver operating curve (AUROC),[25] categorization into three groups (≤24, 25-29 and ≥30

years), and categorization into five groups (≤24, 25-29,30-34, 35-39 and ≥40 years). The model with age categorization that yielded the lowest Akaike Information Criterion (AIC), highest log likelihood, pseudo R2, and AUROC was selected for inclusion in the risk score.

To generate the risk score, the number of points assigned for each included variable was determined by dividing each logistic regression coefficient in the final multivariable model by the value of the smallest regression coefficient, then rounding to the nearest integer. Sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) were calculated for different risk score thresholds to determine the optimal threshold for the risk score. Using the NPV and PPV, the predictive summary index (PSI) was calculated, using the formula PPV+NPV-1. The numbers needed to screen (NNS) using sensitive NAAT testing to detect one case of CT infection were calculated as 1/PSI for each cutoff.[26] Lastly, two simplified risk scores were generated. The first excluded BV diagnosed by Nugent score from the model since it requires laboratory testing. The second simplified risk score excluded BV and also included a simplified assessment of alcohol intake; in place of the full AUDIT score, the first question of the AUDIT interview was used ("*How often do you have a drink containing alcohol, such as beer, wine, spirits, or locally brewed alcohol (such as "changaa", "manzi", "bus", or "muratina?")*.

Several model characteristics were assessed. To assess model performance, the Brier score was determined. The Brier score ranges from zero where the model is perfect to 0.25 for a model that is not useful.[27,28] To assess model calibration, a calibration plot for observed and predicted outcomes grouped in deciles was generated. Lastly, model discrimination was based on the c-statistic which ranges from 0.5 (no discrimination) to 1 (perfect discrimination).[27]

RESULTS

Of 703 participants enrolled in the MPTB study, 691 (98.2%) had STI testing by NAAT results at enrollment, and were included in this analysis. Almost all participants were Kenyan (687/691, 99.4%), Protestant or Catholic (632/690, 91.6%), and married (665/691, 96.2%) (Table1). The median age of participants was 29 years (interquartile range, IQR: 25-34). Sixty-two participants (62/690, 9.0%) had partners who were older by more than 10 years. Regarding education status, 176/691 (25.5%) had achieved primary education or lower, 98/691 (14.2%) had some secondary education, 302/691 (43.7%) completed secondary school or had some higher education, and 115/691 (16.6) had completed college education. Two-thirds of participants (446/686, 65.1%) had a monthly household income above 10,000 Kenya shillings (~97 USD-2020).

By design, no participants were on any form of contraception at the time of enrollment. However, some participants (393/688, 57.1%) reported contraception in the month prior to enrollment including implants, (86/688, 27%), copper IUD (132/688, 19.1%), condoms (42/688, 6.1%), depo medroxyprogesterone acetate (DMPA) (21/688, 3.0%) and natural methods (6/691, 0.9%). Most participants (623/683, 91.4%) reported having had sex without a condom in the past month. Six hundred and seven (607/691, 87.8%) participants had a score of zero on the AUDIT score, 72/691 (10.4%) had scores between 1 and 7, while only 12 participants (12/691, 1.7%) had an AUDIT score ≥8, suggesting hazardous or harmful drinking.

Few participants reported genitourinary symptoms; 67/691 (9.7%) reported having abnormal vaginal discharge, 66/691 (9.6%) had vulvovaginal itch, and 73/691 (10.6%) reported having abdominal pain. On examination, 9/691 (1.3%) exhibited cervical motion tenderness while 22/691 (3.2%) had clinically detected cervical inflammation. One hundred and nineteen participants (119/691, 17.2%) reported having had a genital infection in the past. The prevalence of BV was 240/691 (34.7%) and the prevalence of vaginal candidiasis was 62/690 (9.0%).

The overall prevalence of any of the three treatable STIs *(C. trachomatis, N. gonorrhoeae, and T. Vaginalis*) was 60/688 (8.7%). *Chlamydia trachomatis* had the highest prevalence at 51/691 (7.4%), while *T. vaginalis* (6/687, 0.9%) and *N. gonorrhoeae* (5/691, 0.7%) were rare. Only two participants had concurrent infections with more than one STI; one had chlamydia and trichomoniasis and the other chlamydia and gonorrhea.

In univariate analysis assessing correlates of any STI, women younger than 25 years (OR 2.41; 95%CI: 0.69-8.48) and those whose partners were younger than 25 years (OR 17.22; 95%CI: 3.74-79.24) had a higher odds of having any of the three STIs (Table 2). Participants who had at least some secondary education (OR 2.61; 95%CI: 1.07-6.39), those with AUDIT score ≥8 (OR 5.54; 95%CI: 1.61-19.04), and those with BV (OR 2.49; 95%CI: 1.46-4.24) also had higher odds of having STIs. There was no association between abnormal vaginal discharge (OR 1.24; 95%CI: 0.54-2.87) or lower abdominal pain (OR 0.75 (95%CI: 0.54-2.87) and STI. In the multivariate analysis, BV (adjusted OR, aOR- 2.30; 95%CI: 1.31-4.04) and having a high AUDIT score (aOR 4.86; 95%CI: 1.30-18.24) remained significantly associated with a higher odds of any STI.

Correlates of prevalent *C. trachomatis* infection included age <25 years (OR 6.75; 95%CI: 0.88-51.68), partners' age <25 (52.33; 95%CI: 5.73, 477.73), BV (OR 2.30; 95%CI: 1.31-4.04), and AUDIT score ≥8 (OR 5.54; 95%CI: 1.61-19.04) (Table 3). Symptoms including self-reported abnormal vaginal discharge (OR 1.01; 95%CI: 0.39-2.64) and lower abdominal pain (OR 0.70 (95%CI: 0.25-2.01) were not associated with *C. trachomatis* infection. In the multivariate analysis, BV (aOR 2.28; 95%CI: 1.25-4.16) and AUDIT score>=8 (aOR 5.71; 95%CI: 1.40-21.69) remained significantly associated with higher odds of *C. trachomatis*.

The variables in the final model for the risk score derivation for predicting *C. trachomatis* infection, based on a p-value <0.10 on univariate analysis, included participants' age (≤24 years, 25-29 years and ≥30 years), AUDIT score (0-7 versus ≥8), and BV (Table 4). This model yielded the lowest AIC and highest pseudo R2 compared to the other models utilizing other age

categorizations. The AUROC for this final model was 0.78 and the Brier score was 0.06. The calibration plot of observed versus expected probability of the occurrence of *C. trachomatis* infection was almost along the 45 degrees line indicating a near-perfect prediction for lower probabilities and over-prediction by the model when the expected probability was higher (Fig 2).

Using the regression coefficients of the logistic regression model, the number of total points for each included predictor variable was calculated. The final risk score included a range of 0-6. The AUROC for the final risk score was 0.78 (95% CI: 0.72-0.84) (Fig.1). The sensitivity, specificity, positive and negative predictive value, AUROC, and NNS for each possible cutoff in the risk score is presented in Table 5. For example, at a cutoff of ≤2 versus ≥3, 220 participants (31.8%) were classified as higher risk for *C. trachomatis* infection. This cutoff yielded a sensitivity of 70.6%% (95%CI: 56.2-82.5) and a specificity of 71.3% (95%CI: 67.7-74.7). At this cutoff, the number needed to screen (NNS) to identify one individual with *C. trachomatis* infection was 7.69.

The first simplified model excluded BV, and included participants' age and AUDIT score per the complete model described above. This first simplified model generated a risk score ranging from 0-3 with AUDIT score>=8 (score=1) and age (≤24 years score=2, 25-29 years score=1 and ≥30 years score=0), and had an AUROC (0.752; 95%CI: 0.691-0.813, Table 4). For the second simplified score, BV was also excluded. The model included a single question about alcohol use (how often a participant drank alcohol (Never or less than one time monthly score=0, ≥2 times a month score=1) and participants' age (24 years score=2, 25-29 years score=1 and ≥30 years score=0). The AUROC for this second simplified score was 0.747 (95% CI: 0.687-0.808). The sensitivity, specificity, PPV, NPV and NNS for these two simplified risk scores are presented in Table 6.

DISCUSSION

This is the first study to report the prevalence and correlates of treatable STIs among Kenyan women who are actively planning a pregnancy. The overall prevalence of *C. trachomatis, N. gonorrhoeae*, and *T. vaginalis* infections was 8.7%, with *C. trachomatis* being the most prevalent (7.4%). None of the symptoms suggestive of an STI, such as vaginal discharge or lower abdominal pain, was significantly associated with the presence of an STI. Women were more likely to have an STI if they were younger, had younger male partners, reported hazardous or harmful alcohol use, or had BV. The final risk score for prediction of *C. trachomatis* infection included participant age, AUDIT score, and BV. At a binary cut point of ≤2 vs. ≥3, 70% of cases were detected in the high risk subset including only 30% of the population. Using the simplified risk score including only age and a single question about alcohol use that may be more easily adaptable to clinical settings in Kenya, a risk score cutoff of ≤1 vs ≥2 identified 22% of the population as higher risk for prevalent *C. trachomatis* and would have detected 55% of the infections.

Identifying pregnancy planners at increased risk for prevalent STIs and screening them using sensitive NAATs could reduce complications of undetected infections including infertility, onward transmission, preterm birth, and neonatal infections.[29] In LMICs, risk scoring tools have the potential to facilitate a more cost-effective approach to screening by identifying higher risk individuals versus screening the entire population.[30] The complete risk score developed to identify Kenyan pregnancy planners at high risk for prevalent *C. trachomatis* had an AUROC of 0.78 (95% Cl: 0.72-0.84) which is comparable to other risk scores for *C. trachomatis* in literature. For example, a risk score using demographic, sexual behavior, and clinical features of heterosexual women attending a sexual health clinic in Australia had an AUROC of 0.72 (95% Cl: 0.70-0.74),[31] while one developed in Canada in a population of asymptomatic sexual health clinic attendees yielded an AUROC of 0.74. (95% Cl: 0.70-0.77).[30] Some authors suggest that a good screening tool is one that can detect 90% of cases by screening less than 60% the population[32].

However, in low-resource settings where no STI testing is typically performed in pregnancy planners, the optimal cutoff should be sensitive to cost and budget impact. Such a cutoff could potentially provide an important individual and population-level benefit even if <90% of cases were captured.

Consistent with the findings of other studies, this study has demonstrated that clinical symptoms such as abnormal vaginal discharge, macroscopic evidence of cervicitis, and lower abdominal pain are not significantly associated with the presence of an STI.[33,34] Currently, most LMICs rely on syndromic management for detection and treatment of STIs.[12,35] Syndromic management not only misses those who are asymptomatic and those without contact with healthcare, but also results in overtreatment of those who are symptomatic.[12,36] The overtreatment posed by syndromic management may introduce antibiotic resistance and further transmission of drug resistant microbes while those who are asymptomatic remain undetected and vulnerable to complications.[12,33] The limitations of syndromic management have led many to advocate for other approaches.[12,33,37,38] Risk scores that reliably identify higher risk women for screening may provide a feasible approach for LMICs.

Epidemiologic patterns of STIs vary by region, but prevalence is high in sub-Saharan Africa.[1] Across several studies in a variety of populations in sub-Saharan Africa, the prevalence of *C. trachomatis* has ranged from 1-31%. Higher prevalences have generally been observed in younger women, sex workers, students, adolescents, and those attending gynecological clinics.[21,39–41] Still, in some studies, women who were classified as low-risk, such those attending ANC (31%) and family planning clinics (18%), had high prevalences of STIs. Studies of *N. gonorrhoeae* in Africa have found prevalences between 2-8%,[39–41] with higher prevalences in sex workers. Finally, the prevalence of *T. vaginalis* in African populations has ranged from 5-23%. Higher prevalences have been observed in HIV-positive women and those with another STI.[26,30-32] Among women attending the family planning clinic at a referral hospital in Kenya, the prevalence

of *C. trachomatis* was 13%, which is nearly double the prevalence in this cohort of pregnancy planners (7.4%).[42]

The risk factors associated with prevalent STIs in this cohort included younger age, presence of BV, and hazardous or harmful alcohol use by AUDIT, which are similar to risk factors presented in other studies.[21,43,44] The present analysis did not find a significant association between partners' age difference and STI risk, a factor that has been associated with STIs in a number of other studies.[45,46]

This study has several unique strengths. First, the study included a cohort that was recruited primarily from family planning clinics as they sought to discontinue their contraceptive method to conceive. Thus, the findings may be applicable to other groups of general-population African women planning to conceive, and might also be extrapolated to the early antenatal period. Second, using statistical methods, this study has derived complete and simplified risk scores with variable cutoffs that have the potential to identify a high proportion of high-risk women with prevalent, treatable STIs during the preconception period.

There were also several limitations. First, the risk scoring tool has not yet been internally validated; this will be performed using bootstrap techniques at a later point. Second, women may have been less likely to report alcohol use due to social desirability bias. Underreporting of alcohol use was likely unrelated to *C. trachomatis* infection status or for any STI, which would result in non-differential misclassification. This would tend to attenuate the observed association between alcohol use and any STI or chlamydia alone. Despite this limitation, this study identified a strong and significant association between hazardous/harmful alcohol use and *C. trachomatis* infection as well as any STI, and this variable served as a useful part of the risk score for chlamydia. Lastly, because this cohort includes women planning pregnancies, the risk score may not be generalizable to women with unintended pregnancies or women considered high-risk, such as those with multiple partners.

In conclusion, the risk-scoring tool developed using data from this cohort of Kenyan women planning pregnancies identified high-risk individuals for *C. trachomatis* screening and would have detected the majority of *C. trachomatis* infections while screening less than half of the study population. This tool will be internally validated using bootstrap resampling and externally validated at a later point. In addition, micro-costing analyses based on implementation of a two-stage screening process using the risk score with NAAT testing for higher-risk women will be helpful for estimating the budget impact of this type of screening program, and the cost per STI treated. Ultimately, this risk-scoring tool could provide a simple and cost-effective approach to reduce the rate of untreated *C. trachomatis* infection in Kenyan pregnancy planners.

REFERENCES

1. Rowley J, Vander Hoorn S, Korenromp E, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bull World Health Organ*. 2019;97(8):548-562P. doi:10.2471/BLT.18.228486

2. World Health Organization. Sexually transmitted infections (STIs). Published 2019. Accessed June 15, 2020. https://www.who.int/news-room/fact-sheets/detail/sexually-transmitted-infections-(stis)

3. Institute of Medicine (US) Committee on Prevention and Control of Sexually Transmitted Diseases. *The Hidden Epidemic: Confronting Sexually Transmitted Diseases: Summary*. (Eng TR, Butler WT, eds.). National Academies Press (US); 1997. Accessed June 15, 2020. http://www.ncbi.nlm.nih.gov/books/NBK233453/

4. Tsevat DG, Wiesenfeld HC, Parks C, Peipert JF. Sexually transmitted diseases and infertility. *American Journal of Obstetrics and Gynecology*. 2017;216(1):1-9. doi:10.1016/j.ajog.2016.08.008

5. Lassi ZS, Imam AM, Dean SV, Bhutta ZA. Preconception care: preventing and treating infections. *Reprod Health*. 2014;11(Suppl 3):S4. doi:10.1186/1742-4755-11-S3-S4

6. Galvin SR, Cohen MS. The role of sexually transmitted diseases in HIV transmission. *Nature Reviews Microbiology*. 2004;2(1):33-42. doi:10.1038/nrmicro794

7. Masson L, Passmore J-AS, Liebenberg LJ, et al. Genital Inflammation and the Risk of HIV Acquisition in Women. *Clin Infect Dis*. 2015;61(2):260-269. doi:10.1093/cid/civ298

8. Johnson LF, Lewis DA. The effect of genital tract infections on HIV-1 shedding in the genital tract: a systematic review and meta-analysis. *Sex Transm Dis*. 2008;35(11):946-959. doi:10.1097/OLQ.0b013e3181812d15

9. Cohen MS, Council OD, Chen JS. Sexually transmitted infections and HIV in the era of antiretroviral treatment and prevention: the biologic basis for epidemiologic synergy. *J Intern AIDS Soc*. 2019;22(S6). doi:10.1002/jia2.25355

10. Workowski, Kimberly A, Bolan, Gail A. Sexually Transmitted Diseases Treatment Guidelines, 2015. Published 2015. Accessed June 15, 2020. https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6403a1.htm

11. Meyers DS, Wolff T, Gregory K, et al. USPSTF Recommendations for STI Screening. *AFP*. 2008;77(6):819-824.

12. Wi TE, Ndowa FJ, Ferreyra C, et al. Diagnosing sexually transmitted infections in resource-constrained settings: challenges and ways forward. *Journal of the International AIDS Society*. 2019;22(S6):e25343. doi:10.1002/jia2.25343

13. WHO. WHO | Guidelines for the management of sexually transmitted infections. WHO. Published 2004. Accessed June 18, 2020. https://www.who.int/hiv/pub/sti/pub6/en/

14. Meade JC, Cornelius DC. Sexually Transmitted Infections in the Tropics. *Current Topics in Tropical Medicine*.:29.

15. National AIDS and STI Control Programme. *Kenya National Guidelines for Prevention, Management and Control of Sexually Transmitted Infections*.; 2018.

16. Torrone EA, Morrison CS, Chen P-L, et al. Prevalence of sexually transmitted infections and bacterial vaginosis among women in sub-Saharan Africa: An individual participant data meta-analysis of 18 HIV prevention studies. *PLoS Med*. 2018;15(2). doi:10.1371/journal.pmed.1002511

17. Chirenje ZM, Gundacker HM, Richardson B, et al. Risk Factors for Incidence of Sexually Transmitted Infections Among Women in a Human Immunodeficiency Virus Chemoprevention Trial: VOICE (MTN-003). *Sex Transm Dis*. 2017;44(3):135-140. doi:10.1097/OLQ.0000000000000568

18. Gaydos CA, Jett-Goheen M, Barnes M, Dize L, Barnes P, Hsieh Y-H. Use of a risk quiz to predict infection for sexually transmitted infections: a retrospective analysis of acceptability and positivity. *Sex Transm Infect*. 2016;92(1):44-48. doi:10.1136/sextrans-2015-052058

19. Ginindza TG, Stefan CD, Tsoka-Gwegweni JM, et al. Prevalence and risk factors associated with sexually transmitted infections (STIs) among women of reproductive age in Swaziland. *Infect Agent Cancer*. 2017;12. doi:10.1186/s13027-017-0140-y

20. Tun W, Sheehy M, Broz D, et al. HIV and STI Prevalence and Injection Behaviors Among People Who Inject Drugs in Nairobi: Results from a 2011 Bio-behavioral Study Using Respondent-Driven Sampling. *AIDS Behav*. 2015;19(Suppl 1):24-35. doi:10.1007/s10461-014-0936-3

21. Hussen S, Wachamo D, Yohannes Z, Tadesse E. Prevalence of chlamydia trachomatis infection among reproductive age women in sub Saharan Africa: a systematic review and meta-analysis. *BMC Infect Dis*. 2018;18(1):596. doi:10.1186/s12879-018-3477-y

22. Lokken EM, Mandaliya K, Srinivasan S, et al. Impact of preconception vaginal microbiota on women's risk of spontaneous preterm birth: protocol for a prospective case-cohort study. *BMJ Open*. 2020;10(2). doi:10.1136/bmjopen-2019-035186

23. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J Clin Microbiol*. 1991;29(2):297-301.

24. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Logistic regression. *Perspect Clin Res*. 2017;8(3):148-151. doi:10.4103/picr.PICR_87_17

25. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem Med (Zagreb)*. 2016; 26(3):297-307. doi:10.11613/BM.2016.034

26. Larner AJ. Number Needed to Diagnose, Predict, or Misdiagnose: Useful Metrics for Non-Canonical Signs of Cognitive Status? *DEE*. 2018;8(3):321-327. doi:10.1159/000492783

27. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*. 2010;21(1):128-138. doi:10.1097/EDE.0b013e3181c30fb2

28. Cook NR. Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction. *Circulation*. 2007;115(7):928-935. doi:10.1161/CIRCULATIONAHA.106.672402

29. Adachi K, Nielsen-Saines K, Klausner JD. Chlamydia trachomatis Infection in Pregnancy: The Global Challenge of Preventing Adverse Pregnancy and Infant Outcomes in Sub-Saharan Africa and Asia. *Biomed Res Int*. 2016;2016. doi:10.1155/2016/9315757

30. Falasinnu T, Gilbert M, Gustafson P, Shoveller J. Deriving and Validating A Risk Estimation Tool for Screening Asymptomatic Chlamydia and Gonorrhea. *Sexually Transmitted Diseases*. 2014;41(12):706–712. doi:10.1097/OLQ.0000000000000205

31. Wand H, Guy R, Donovan B, McNulty A. Developing and validating a risk scoring tool for chlamydia infection among sexual health clinic attendees in Australia: a simple algorithm to identify those at high risk of chlamydia infection. *BMJ Open*. 2011;1(1). doi:10.1136/bmjopen-2010-000005

32. La Montagne DS, Patrick LE, Fine DN, Marrazzo JM, Project OB of the RXIP. Re-evaluating Selective Screening Criteria for Chlamydial Infection Among Women in the U.S. Pacific Northwest. *Sexually Transmitted Diseases*. 2004;31(5):283–289. doi:10.1097/01.OLQ.0000124613.85111.6B

33. Sonkar SC, Wasnik K, Kumar A, et al. Evaluating the utility of syndromic case management for three sexually transmitted infections in women visiting hospitals in Delhi, India. *Sci Rep*. 2017;7. doi:10.1038/s41598-017-01422-y

34. Dela H, Attram N, Behene E, et al. Risk factors associated with gonorrhea and chlamydia transmission in selected health facilities in Ghana. *BMC Infectious Diseases*. 2019;19(1):425. doi:10.1186/s12879-019-4035-y

35. Ghebremichael M. The Syndromic versus Laboratory Diagnosis of Sexually Transmitted Infections in Resource-Limited Settings. *ISRN AIDS*. 2014;2014. doi:10.1155/2014/103452

36. Cristillo AD, Bristow CC, Peeling R, et al. Point-of-Care Sexually Transmitted Infection Diagnostics: Proceedings of the STAR Sexually Transmitted Infection—Clinical Trial Group Programmatic Meeting. *Sex Transm Dis*. 2017;44(4):211-218. doi:10.1097/OLQ.0000000000000572

37. Kularatne RS, Niit R, Rowley J, et al. Adult gonorrhea, chlamydia and syphilis prevalence, incidence, treatment and syndromic case reporting in South Africa: Estimates using the Spectrum-STI model, 1990-2017. *PLOS ONE*. 2018;13(10):e0205863. doi:10.1371/journal.pone.0205863

38. Gupta. Syndromic management of sexually transmitted infections: A critical appraisal and the road ahead. Accessed August 6, 2020. http://www.nmji.in/article.asp?issn=0970-258X;year=2019;volume=32;issue=3;spage=147;epage=152;aulast=Gupta

39. Menéndez C, Castellsagué X, Renom M, et al. Prevalence and Risk Factors of Sexually Transmitted Infections and Cervical Neoplasia in Women from a Rural Area of Southern Mozambique. Infectious Diseases in Obstetrics and Gynecology. doi:https://doi.org/10.1155/2010/609315

40. Ahmed K, Baeten JM, Beksinska M, et al. HIV incidence among women using intramuscular depot medroxyprogesterone acetate, a copper intrauterine device, or a levonorgestrel implant for contraception: a randomised, multicentre, open-label trial. *The Lancet*. 2019;394(10195):303-313. doi:10.1016/S0140-6736(19)31288-7

41. Dubbink JH, Verweij SP, Struthers HE, et al. Genital Chlamydia trachomatis and Neisseria gonorrhoeae infections among women in sub-Saharan Africa: A structured review: *International Journal of STD & AIDS*. Published online February 28, 2018. doi:10.1177/0956462418758224

42. Maina AN, Kimani J, Anzala O. Prevalence and risk factors of three curable sexually transmitted infections among women in Nairobi, Kenya. *BMC Res Notes*. 2016;9(1):193. doi:10.1186/s13104-016-1990-x

43. Bautista CT, Wurapa EK, Sateren WB, Morris SM, Hollingsworth BP, Sanchez JL. Association of Bacterial Vaginosis With Chlamydia and Gonorrhea Among Women in the U.S. Army. *American Journal of Preventive Medicine*. 2017;52(5):632-639. doi:10.1016/j.amepre.2016.09.016

44. Bautista CT, Wurapa E, Sateren WB, Morris S, Hollingsworth B, Sanchez JL. Bacterial vaginosis: a synthesis of the literature on etiology, prevalence, risk factors, and relationship with chlamydia and gonorrhea infections. *Military Medical Research*. 2016;3(1):4. doi:10.1186/s40779-016-0074-5

45. Kelly RJ, Gray RH, Sewankambo NK, et al. Age differences in sexual partners and risk of HIV-1 infection in rural Uganda. *J Acquir Immune Defic Syndr*. 2003;32(4):446-451. doi:10.1097/00126334-200304010-00016

46. Maughan-Brown B, Kenyon C, Lurie MN. Partner age differences and concurrency in South Africa: Implications for HIV-infection risk among young women. *AIDS Behav*. 2014;18(12):2469-2476. doi:10.1007/s10461-014-0828-6

47. Staras SAS, Cook RL, Clark DB. Sexual Partner Characteristics and Sexually Transmitted Diseases Among Adolescents and Young Adults. *Sex Transm Dis*. 2009;36(4). doi:10.1097/OLQ.0b013e3181901e32

48. Longfield K, Glick A, Waithaka M, Berman J. Relationships Between Older Men and Younger Women: Implications for STIs/HIV in Kenya. *Studies in Family Planning*. 2004;35(2):125-134. doi:10.1111/j.1728-4465.2004.00014.x

49. Beauclair R, Kassanjee R, Temmerman M, Welte A, Delva W. Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa. *The European Journal of Contraception & Reproductive Health Care*. 2012;17(1):30-39. doi:10.3109/13625187.2011.644841

Table 1: Characteristics of the 691 MPTB study participants at enrollment

Characteristic	N	n(%) or median (IQR)
Demographic & Behavioral and Reproductive Characteristics		
Age (years)	691	29 (25,34)
Age categories (years)	691	
≤24		147 (21.3)
25-29		222 (32.1)
30-34		172 (24.9)
35-39		119 (17.2)
≥40		31 (4.5)
Partner's age in categories (years)	690	
≤24		20 (2.9)
25-29		138 (20.0)
30-34		196 (28.4)
35-39		178 (25.8)
≥40		158 (22.9)
Nationality (Kenyan)	691	687 (99.4)
Education	691	
None or some/completed primary (zero to 8 years)		176 (25.5)
Some secondary (9 to 11 years)		98 (14.2)
Completed secondary or some higher education (12 to 15 years)		302 (43.7)
Completed college (≥16 years)		115 (16.6)
Religion	690	
Protestant		482 (69.9)
Catholic		150 (21.7)

Muslim		45 (6.5)
Other		13 (1.8)
Monthly household income (Kenyan Shillings)	686	
<2,500		20 (2.9)
2,500-10,000		220 (32.1)
10,000-30,000		297 (43.3)
30,000-75,000		104 (15.2)
>75,000		45 (6.6)
Marital Status	691	
Married or living together		665 (96.2)
Separated or divorced		19 (2.8)
Never married		5 (0.7)
Widow		2 (0.3)
Age difference with partner[1]	690	
<5 years		369 (53.5)
5-10 years		259 (37.5)
≥10 years		62 (9.0)
Partner HIV status as reported by participant	689	
Negative		545 (79.1)
Positive		33 (4.8)
Unsure		111 (16.1)
Alcohol Use Disorders Identification Test (AUDIT)	691	
AUDIT score 0 (No alcohol use)		607 (87.8)
AUDIT score 1-7 (Minimal risk)		72 (10.4)
AUDIT score ≥8 (Hazardous or harmful alcohol use)		12 (1.7)
Recent contraceptive use [2]	688	

None		295 (42.9)
Condoms only		42 (6.1)
OCPs		6 (0.9)
Injectable (DMPA)		21 (3.0)
Copper IUD		132 (19.1)
Implant		186 (27.0)
Natural methods		6 (0.9)
Self-reported sex without a condom in the past month	683	
0 times		59 (8.6)
1-4 times		211 (30.9)
5-8 times		164 (24.0)
≥9 times		249 (36.5)
Vaginal washing in the past month	691	260 (37.6)
Gravidity and parity	690	
Never been pregnant		74 (10.7)
Pregnant, no live birth (ectopic/abortion)		40 (5.8)
≥1 live birth		576 (83.5)
Clinical characteristics		
Participants report of lower abdominal pain	691	73 (10.6)
Participants report of abnormal vaginal discharge	691	67 (9.7)
Participants report of vulvovaginal itch	691	66 (9.6)
Participants report of history of genital infection[3]	691	119 (17.2)
Vaginal discharge reported by clinician	691	
None/small		437 (63.2)
Moderate/profuse		248 (35.9)
Menses		6 (0.9)

Clinically detected cervical mucopus (inflammation)	691	22 (3.2)
Cervical motion tenderness	691	9 (1.3)
Laboratory characteristics		
Nugent score	691	
Negative for BV (Nugent <7)		451 (65.3)
Positive for BV (Nugent ≥7)		240 (34.7)
Vaginal candidiasis	690	62 (9.0)
Chlamydia trachomatis	691	51 (7.4)
Neisseria gonorrhoeae	691	5 (0.7)
Trichomonas vaginalis	687	6 (0.9)
Prevalence of any STI [4]	688	60 (8.7)

[1] 97.2% of the study participants were younger than their spouses while 2.8% were older. Three (0.43%) participants were older than their spouses by more than 5 years. Participants who were older than their partners were classified as "Less than 5 years category" for this study, as only male partner age greater than female partner's age has been associated with STI risk).[48,49]

[2]Recent contraceptive use was defined as use of OCP, copper IUD, or implant within the past month and last DMPA injection within the past 6 months.

[3]History of treatment for pelvic inflammatory disease, genital ulcer, syphilis, chlamydia, gonorrhoea trichomoniasis, cervicitis, bacterial vaginosis, or STI treatment given empirically.

[4]Laboratory diagnosis by NAAT for Chlamydia trachomatis, Neisseria gonorrhoeae, and Trichomonas vaginalis

Table 2: **Correlates of *Chlamydia trachomatis, Neisseria gonorrhoeae, and / or Trichomonas vaginalis* infection**

Characteristic (N=688)	**Proportion of each category N (%)**	**Prevalence of any STI n (%)**	**Univariate analysis OR (95% CI)**	**Joint p value**	**Multivariate analysis (Any STI*) OR (95% CI)**
Age categories					
≤24	146 (21.2)	30 (20.6)	2.41 (0.69,8.48)	<0.001	2.58 (0.71,9.39)
25-29	221 (32.1)	20 (9.1)	0.93 (0.26,3.33)		1.15 (0.31,4.22)
30-34	171 (24.9)	6 (3.5)	0.34 (0.08,1.44)		0.38 (0.09,1.64)
35-39	119 (17.3)	1 (0.8)	0.08 (0.01,0.79)		0.10 (0.01,1.00)
≥40	31 (4.5)	3 (9.7)	1.00 (ref)		1.00 (ref)
Partner's age in categories (years) [1]					
≤24	20 (2.9)	5 (25.0)	17.22 (3.74,79.24)	<0.001	-
25-29	137 (19.9)	22 (16.1)	9.88 (2.89,33.82)		-
30-34	195 (28.4)	23 (11.8)	6.91 (2.03,23.46)		-
35-39	177 (25.8)	7 (3.6)	2.12 (0.54,8.37)		-
≥40	158 (23.0)	3 (1.9)	1.00 (ref)		-
Education					
None or some/completed primary, 1-8 years	175 (25.4)	14 (8)	1.16 (0.47, 2.87)	0.067	0.77 (0.29,2.03)
Some secondary, 9-11 years	98 (14.2)	16 (16.3)	2.61 (1.07,6.39)		1.47 (0.55,3.93)

Completed secondary or some tertiary education, ≥12 & <16 years	300 (43.6)	22 (7.3)	1.06 (0.46,2.45)		0.69 (0.28,1.70)
Completed college, ≥16 years	115 (16.7)	8 (7.0)	1.00 (ref)		1.00 (ref)
Household Income in Kenyan Shillings					
<2,500	45 (6.6)	3 (6.7)	1.56 (0.24,10.11)	0.246	-
2,500-10,000	104 (15.2)	4 (3.9)	1.72 (0.50,5.99)		-
10,000-30,000	295 (43.1)	26 (8.8)	1.35 (0.39,4.67)		-
30,000-75,000	219 (32.1)	24 (11.0)	0.56 (0.12,2.61)		-
>75,000	20 (2.9)	2 (10.0)	1.00 (ref)		-
Age difference with partner					
<5 years	371 (54.0)	33 (8.9)	1.00 (ref)	0.482	-
5-10 years	255 (37.1)	24 (9.4)	1.06 (0.61,1.85)		-
≥10 years	61 (8.9)	3 (4.9)	0.53 (0.16,1.78)		-
Partner HIV status as reported by participant					
Negative	542 (79.0)	48 (8.7)	1.00 (ref)	0.677	-
Positive	33 (4.8)	4 (12.1)	1.42 (0.48,4.21)		-
Unsure	111 (16.2)	8 (7.2)	0.80 (0.37,1.74)		-
Alcohol Use Disorders Identification Test (AUDIT)					
AUDIT score 0 (No alcohol use)	604 (87.8)	50 (8.3)	1.00 (ref)	0.053	1.00 (ref)

AUDIT score 1-7 (Minimal risk)	72 (10.5)	6 (8.8)	1.01 (0.42,2.44)		1.18 (0.46,3.03)
AUDIT score ≥8 (Hazardous or harmful alcohol use)	12 (1.7)	4 (33.3)	5.54 (1.61,19.04)		4.86 (1.30,18.24)
Recent contraceptive use					
None	293 (42.8)	31 (10.6)	1.00 (ref)	0.219	-
Condoms only	42 (6.1)	2 (4.8)	0.42 (0.10,1.83)		-
OCPs	6 (0.9)	1 (16.8)	1.69 (0.19,14.9)		-
Injectable (DMPA)	21 (3.1)	1 (4.8)	0.42 (0.05,3.26)		-
Copper IUD	132 (19.3)	6 (4.6)	0.40 (0.16,0.99)		-
Implant	185 (27.0)	19 (10.3)	0.97 (0.53,1.77)		-
Natural methods	6 (0.9)	0(0)	-		-
Self-reported frequency of sex without a condom in the past month					
0 times	58 (8.5)	3 (5.2)	1.00 (ref)	0.707	-
1-4 times	211 (31.0)	18 (8.5)	1.71 (0.49,6.02)		-
5-8 times	163 (24.0)	16 (9.8)	2.00 (0.56,7.12)		-
≥9 times	248 (36.5)	23 (9.3)	1.87 (0.54,6.47)		-
Vaginal washing					
No	430 (62.5)	33 (7.7)	1.00 (ref)	0.213	-
Yes	258 (37.5)	27 (10.5)	1.41 (0.82,2.40)		-
Actively trying to conceive prior to enrollment into the study					

No	404 (58.7)	33 (8.1)	1.00 (ref)	0.541	-
Yes	284 (41.3)	27 (9.5)	1.18 (0.69,2.01)		-
Gravidity and parity					
Never been pregnant	73 (10.6)	5 (6.9)	1.00 (ref)	0.362	-
Pregnant, no live birth (ectopic/abortion)	40 (5.8)	6 (15.0)	2.40 (0.68,8.43)		-
≥1 live birth	574 (83.6)	49 (8.5)	1.27 (0.49, 3.30)		-
Participants report of lower abdominal pain					
No	615 (83.4)	55 (8.9)	1.00 (ref)	0.536	-
Yes	73 (10.6)	5 (6.8)	0.75 (0.29,1.93)		-
Participants report of abnormal vaginal discharge					
No	621 (90.3)	53 (8.5)	1.00 (ref)	0.607	-
Yes	67 (9.7)	7 (10.5)	1.25 (0.54, 2.87)		-
Participants report of vulvovaginal itch					
No	623 (90.6)	55 (8.8)	1.00 (ref)	0.753	-
Yes	65 (9.4)	5 (7.7)	0.86 (0.33, 2.23)		-
Participants report of history of an STI					
None	570 (82.9)	51 (9.0)	1.00 (ref)	0.638	-
Yes	118 (17.2)	9 (7.6)	0.84 (0.40,1.76)		-

Vaginal discharge reported by clinician					
None/small	434 (63.1)	36 (8.3)	1.00 (ref)	0.542	-
Moderate/profuse	248 (36.1)	24 (9.7)	1.18 (0.69,2.04)		-
Menses	6 (0.9)	0 (0)	-		-
Clinically detected cervical mucopus, inflammation					
None	666 (96.8)	57 (8.6)	1.00 (ref)	0.438	-
Present	22 (3.2)	3 (13.6)	1.69 (0.48,5.87)		-
Bacterial vaginosis, BV					
Negative for BV, Nugent <7	448 (65.1)	27 (6.0)	1.00 (ref)	<0.001	1.00 (ref)
Positive for BV, Nugent ≥7	240 (34.9)	33 (13.8)	2.49 (1.46,4.24)		2.30 (1.31,4.04)

[1]Partner's age yielded a low p-value and high odds ratios in univariate analysis; however, it was excluded from the multivariate model due to collinearity with participants' age. The two continuous variables yielded a correlation coefficient of 0.78 in bivariate correlation analysis. Participants' age only was therefore included in the multivariate analysis.

Table 3: Correlates of *Chlamydia trachomatis* infection

Characteristic (N=691)	Proportion of each category N (%)	Prevalence n (%)	Univariate analysis OR (95% CI)	Joint p value	Multivariate analysis (Any STI*) OR (95% CI)
Age categories					
≤24	147 (21.3)	27 (18.3)	6.75 (0.88,51.68)	<0.001	6.81 (0.88,52.81)
25-29	222 (32.1)	18 (8.1)	2.65 (0.34,20.56)		2.96 (0.38,23.23)
30-34	172 (24.9)	4 (2.3)	0.34 (0.08,6.61)		0.74 (0.08,6.95)
35-39	119 (17.2)	1 (0.8)	0.08 (0.02,4.18)		0.29 (0.02,4.78)
≥40	31 (4.5)	1 (3.2)	1.00 (ref)		1.00 (ref)
Partner's age in categories (years)[1]					
≤24	20 (2.9)	5 (25.0)	52.33 (5.73, 477.73)	<0.001	-
25-29	138 (20.0)	18 (13.0)	23.55 (3.10, 178.89)		-
30-34	196 (28.4)	21 (10.7)	18.84 (2.51, 141.68)		-
35-39	178 (25.8)	6 (3.4)	5.48 (0.65,46.0)		-
≥40	158 (22.9)	1 (0.6)	1.00 (ref)		-
Education					
None or some/completed primary, 1-8 years	176 (25.5)	13 (7.4)	1.23 (0.48, 3.18)	0.492	-
Some secondary, 9-11 years	98 (14.2)	11 (11.2)	1.95 (0.73,5.24)		-

Completed secondary or some tertiary education, ≥12 & <16 years	302 (43.7)	20 (6.6)	1.09 (0.45,2.66)		-
Completed college, ≥16 years	115 (16.6)	7 (6.1)	1.00 (ref)		-
Household Income in Kenyan Shillings					
<2,500	45 (6.6)	3 (6.7)	1.56 (0.24,10.11)	0.175	-
2,500-10,000	104 (15.2)	3 (2.9)	1.56 (0.45, 5.44)		-
10,000-30,000	297 (43.3)	20 (6.7)	1.01 (0.29,3.55)		-
30,000-75,000	220 (32.1)	22 (10.0)	0.42 (0.08,2.14)		-
>75,000	20 (2.9)	2 (10.0)	1.00 (ref)		-
Age difference with partner					
<5 years	372 (53.9)	27 (7.3)	1.00 (ref)	0.656	-
5-10 years	257 (37.3)	21 (8.2)	1.13 (0.63,2.06)		-
≥10 years	61 (8.8)	3 (4.9)	0.66 (0.19,2.25)		-
Partner HIV status as reported by participant					
Negative	545 (79.1)	42 (7.7)	1.00 (ref)	0.635	-
Positive	33 (4.8)	3 (9.1)	1.20 (0.35,4.09)		-
Unsure	111 (16.1)	6 (5.4)	0.68 (0.28,1.65)		-
Alcohol Use Disorders Identification Test (AUDIT)					
AUDIT score 0 (No alcohol use)	607 (87.8)	42 (6.9)	1.00 (ref)	0.030	1.00 (ref)

AUDIT score 1-7 (Minimal risk)	72 (10.4)	5 (6.9)	1.00 (0.38,2.62)		1.16 (0.42,3.17)
AUDIT score ≥8 (Hazardous or harmful alcohol use)	12 (1.7)	4 (33.3)	6.73 (1.95,23.3)		5.71 (1.50,21.69)
Recent Contraceptive use					
None	295 (42.9)	27 (9.2)	1.00 (ref)	0.203	-
Condoms only	42 (6.1)	1 (2.4)	0.24 (0.03,1.83)		-
OCPs	6 (0.9)	1 (16.7)	1.99 (0.22,17.6)		-
Injectable (DMPA)	21 (3.1)	0 (0)	-		-
Copper IUD	132 (19.2)	6 (4.6)	0.47 (0.19,1.17)		-
Implant	186 (27.0)	16 (8.6)	0.93 (0.49, 1.79)		-
Natural methods	6 (0.9)	0 (0)	-		-
Self-reported frequency of sex without a condom in the past month					
0 times	59 (8.6)	3 (5.1)	1.00 (ref)	0.662	-
1-4 times	211 (30.9)	13 (6.2)	1.23 (0.34,4.45)		-
5-8 times	164 (24.0)	14 (8.5)	1.74 (0.50, 5.97)		-
≥9 times	249 (36.5)	21 (8.4)	1.72 (0.50,5.97)		-
Vaginal washing					
No	431 (62.4)	28 (6.5)	1.00 (ref)	0.257	-
Yes	260 (37.6)	23 (8.9)	1.40 (0.79,2.48)		-

Actively trying to conceive prior to enrollment into the study					
No	405 (58.6)	30 (7.4)	1.00 (ref)	0.974	-
Yes	285 (41.4)	21 (7.3)	0.99 (0.55,1.77)		-
Gravidity and parity					
Never been pregnant	74 (10.7)	4 (5.4)	1.00 (ref)	0.415	-
Pregnant, no live birth (ectopic/abortion)	40 (5.8)	5 (12.5)	2.50 (0.63,9.90)		-
≥1 live birth	576 (83.5)	42 (7.3)	1.37 (0.48,3.95)		-
Participants report of lower abdominal pain					
No	618 (89.4)	47(7.6)	1.00 (ref)	0.495	-
Yes	73 (10.6)	4 (5.5)	0.70 (0.25,2.01)		-
Participants report of abnormal vaginal discharge					
No	624 (90.3)	45 (7.4)	1.00 (ref)	0.979	-
Yes	67 (9.7)	5 (7.5)	1.01 (0.39, 2.64)		-
Participants report of vulvovaginal itch					
No	625 (90.4)	48 (7.7)	1.00 (ref)	0.324	-
Yes	66 (9.6)	3 (4.6)	0.57 (0.17, 1.89)		-
Participants report of history of an STI					

None	572 (82.8)	43 (7.5)	1.00 (ref)	0.760	-
Yes	119 (17.2)	8 (6.7)	0.89 (0.41,1.94)		-
Vaginal discharge reported by clinician					
None/small	437 (63.2)	31 (7.1)	1.00 (ref)	0.644	-
Moderate/profuse	248 (35.9)	20 (8.1)	1.15 (0.64,2.06)		-
Menses	6 (0.9)	0 (0)	-		-
Clinically detected cervical mucopus, inflammation					
None	669 (96.8)	48 (7.2)	1.00 (ref)	0.302	-
Present	22 (3.2)	3 (13.6)	2.04 (0.58,7.15)		-
Bacterial vaginosis, BV					
Negative for BV, Nugent <7	451 (65.3)	23 (5.1)	1.00 (ref)	0.002	1.00 (ref)
Positive for BV, Nugent ≥7	240 (34.7)	28 (11.7)	2.46 (1.38,4.37)		2.28 (1.25,4.16)

[1]Partner's age yielded a low p-value and high odds ratios in univariate analysis; however, it was excluded from the multivariate model due to collinearity with participants' age. The two continuous variables yielded a correlation coefficient of 0.78 in bivariate correlation analysis. Participants' age only was therefore included in the multivariate analysis

Table 4: **Derivation of the risk scores using logistic regression coefficients for the final models**

	Complete Score				Simplified Score 1				Simplified Score 2			
Variable	**OR (95% CI)**	**P value**	**Logistic regression coefficient**	**Risk score**	**OR (95% CI)**	**P value**	**Logistic regression coefficient**	**Risk score**	**OR (95% CI)**	**P value**	**Logistic regression coefficient**	**Risk score**
Age categories (years)												
≤24	11.11 (4.44,27.82)	<0.001	2.41	3	11.46 (4.60,28.53)	<0.001	2.44	2	11.75 (4.72,29.22)	<0.001	2.46	2
25 to 29	4.90 (1.89,12.67)	0.001	1.59	2	4.68 (1.82,12.01)	0.001	1.54	1	4.55 (1.77,11.68)	0.002	1.52	1
≥30	1.00 (ref)	-	0 (ref)	0	1.00 (ref)	-	0 (ref)	0	1.00 (ref)	-	0 (ref)	0
Bacterial vaginosis												
Negative	1.00 (ref)	-	0 (ref)	0								
Positive	2.30 (1.26,4.20)	0.006	0.84	1								
Alcohol intake (AUDIT[1] Score)												
Low risk (AUDIT score 0-7)	1.00 (ref)	-	0 (ref)	0	1.00 (ref)	-	0 (ref)	0				
High risk (AUDIT score ≥8)	5.62 (1.49, 21.17)	0.011	1.73	2	5.57 (1.44,21.46)	0.013	1.72	1				
Alcohol intake (How often do you drink alcohol?)												
Never or less than one time monthly									1.00 (ref)	-	0 (ref)	0
≥2 times a month									3.10 (0.93,10.30)	0.065	1.13	1
Total Possible Score				6	**Total Possible Score**			**3**	**Total Possible Score**			**3**

**Final model for the complete score: Logit p (Chlamydia infection) = βo (_cons) + β1BV $_{(vaginosis)}$ + β2A$_{Age1}$+ β3A$_{Age2}$ + β4AU$_{Audit}$*

**Final model for the simplified score 1: Logit p (Chlamydia infection) = βo (_cons) + β2A$_{Age1}$+ β3A$_{Age2}$ + β4AU$_{Audit}$*

**Final model for the simplified score 2: Logit p (Chlamydia infection) = βo (_cons) + β2A$_{Age}$1+ β3A$_{Age2}$ + β4$_{Alcohol}$*

[1]AUDIT=Alcohol Use Disorders Identification Test

Table 5: Operating characteristics for binary cutoff points for the full risk score

Score	Low Risk, n (%)	Higher Risk n (%)	Prevalence of chlamydia in high risk group n (%)	Sensitivity % (95%CI)	Specificity % (95%CI)	Positive predictive value % (95%CI)	Negative Predictive value % (95%CI)	No. needed to screen/ predict	Correctly classified-accuracy % (95%CI)	AUROC (95%CI)
0vs≥1	213 (30.82)	478 (69.18)	50 (10.46)	98.04 (89.55,99.95)	33.13 (29.49,36.92)	10.49 (9.87,11.13)	99.53 (96.80,99.93)	9.98	37.92 (34.39,41.66)	0.66 (0.63,0.68)
≤1vs≥2	318 (46.02)	373 (53.98)	46 (12.33)	90.02 (78.59,96.74)	48.91 (44.97,52.85)	12.36 (11.14,13.70)	98.42 (96.44,99.31)	9.27	51.95 (41.96,55.74)	0.68 (0.65,0.74)
≤2vs≥3	471 (68.16)	220 (31.84)	36 (16.36)	70.59 (56.17,82.51)	71.25 (67.67,74.73)	16.40 (13.66,19.57)	96.81 (95.18,97.90)	7.69	71.20 (67.67,74.55)	0.71 (0.64,0.77)
≤3vs≥4	621 (89.87)	70 (10.13)	18 (25.71)	35.29 (22.43,49.93)	91.88 (89.48,93.87)	25.77 (18.07,35.34)	94.67 (93.54,95.61)	4.76	87.70 (85.00,90.05)	0.64 (0.57,0.70)
≤4vs≥5	684 (98.99)	7 (1.01)	2 (28.57)	3.92 (0.48,13.46)	99.22 (98.19,99.75)	28.63 (7.39,66.85)	92.82 (92.44,93.18)	4.55	92.19 (89.91,94.06)	0.52 (0.49,0.54)

Table 6: Operating characteristics estimated for the different binary cutoff points for the simplified scores

Score	Low Risk, n (%)	High Risk (Fraction screened) n (%)	Prevalence of chlamydia in the high risk group n (%)	Sensitivity in % (95%CI)	Specificity in % (95%CI)	Positive predictive value (95%CI)	Negative Predictive value (95%CI)	No. needed to screen/ predict	Correctly classified-accuracy in % (95%CI)	AUROC (95%CI)
Cutoffs for the simplified score with age and assessment of alcohol done using AUDIT										
0 vs≥1	318 (46.02)	373 (53.98)	46 (12.33)	90.20 (78.59,96.74)	48.91 (44.97,52.85)	12.36 (11.14,13.70)	98.42 (96.44,99.31)	9.27	51.95 (48.16,55.74)	0.70 (0.65,0.74)
≤1vs≥2	541 (78.29)	150 (21.71)	28 (18.67)	54.90 (40.34,68.87)	80.94 (77.68,83.91)	18.71 (14.62,23.62)	95.74 (94.30,96.82)	6.67	79.02 (75.78,81.99)	0.68 (0.60,0.75)
≤2vs≥3	686 (99.28)	5 (0.72)	2 (40.00)	3.92 (0.48,13.46)	99.53 (98.64,99.90)	40.07 (10.26,79.64)	92.84 (92.46,93.20)	3.03	92.47 (90.23,94.31)	0.52 (0.49,0.54)
Cutoffs for the simplified score with age and assessment of alcohol done using the first question of AUDIT.										
0 vs≥1	315 (45.59)	376 (54.41)	46 (12.23)	90.20 (78.59,96.74)	48.44 (44.50,52.39)	12.26 (11.05,13.59)	98.41 (96.40,99.30)	9.37	51.52 (47.73,55.51)	0.69 (0.65,0.74)
≤1vs≥2	536 (77.57)	155 (22.43)	28 (18.06)	54.90 (40.34,68.87)	80.16 (76.85,83.18)	18.11 (14.15,22.87)	95.70 (94.25,96.79)	7.24	78.29 (75.02,81.31)	0.68 (0.60,0.75)
≤2vs≥3	686 (99.28)	5 (0.72)	2 (40.00)	3.92 (0.48,13.46)	99.53 (98.64,99.90)	40.07 (10.26,79.64)	92.84 (92.46,93.20)	3.03	92.47 (90.23,94.31)	0.52 (0.49,0.54)

Figure 1: Area under receiver operating curve (AUROC) for the complete risk score and two simplified risk scores

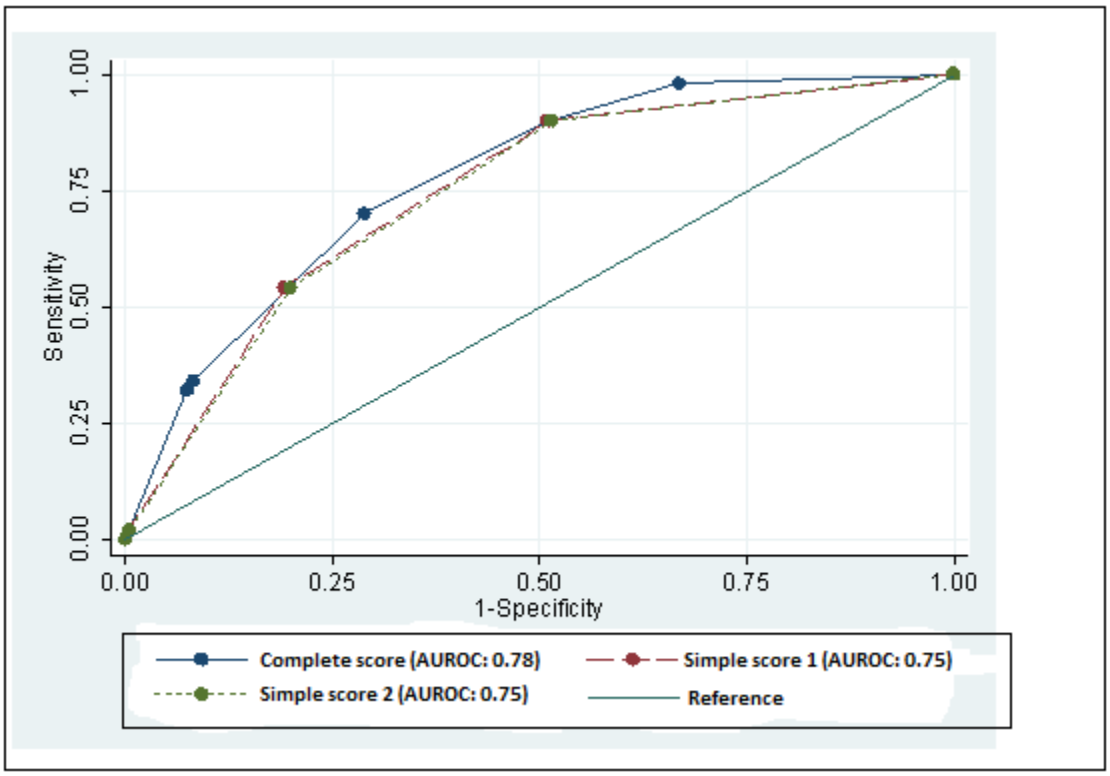

Figure 2: Decile Calibration plot: probability of the actual outcome versus the expected