

Dissertation Proposal: Molly Rosenberg
School Enrollment, Alcohol Outlets, and Sexual Risk

**The Influence of School Enrollment, Alcohol Outlets, and Alcohol Outlet Density on Sexual
Risk among Rural South African Young Women**

**Molly S. Rosenberg, MPH
Dissertation Proposal
University of North Carolina School of Public Health
Department of Epidemiology
May 14, 2013**

**Dissertation committee:
Audrey Pettifor, MPH, PhD (Chair)
Michael Emch, MA, PhD
William Miller, MD, PhD
Harsha Thirumurthy, MA, M.Phil., PhD
Annelies Van Rie, MD, PhD**

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

A. SPECIFIC AIMS

A.1 Overview

The proposed study is designed to examine how and where young South African women spend their time and how these routine activity patterns may influence their sexual risk. We assert that spending time in structured and supervised places will be associated with safer behaviors, either by reducing opportunities to engage in unsafe behaviors or by putting youth in contact with safer friends and partners. Similarly, unstructured and unsupervised activity may provide opportunities to engage in unsafe behaviors or put youth in contact with unsafe friends and partners. An important structured place where many young women may spend time is school. To explore the influence of school enrollment on sexual behavior, using a longitudinal dataset from rural South Africa, we aim to compare pregnancy rates at times when young women are enrolled in school to pregnancy rates at times when young women are not enrolled in school, controlling for appropriate confounders.

Participation in activities aside from school could also influence sexual risk behavior in this population. Specifically, visits to alcohol vendors could allow for periods of unsupervised and unstructured time and provide environments where alcohol consumption and sexual risk behaviors may be promoted. To explore the association between frequency of visits to alcohol vendors and sexual risk behavior, we are proposing a cross-sectional ancillary study to HPTN 068: *The Effects of Cash Transfer and Community Mobilization in Young South African Women Study* in rural South Africa.

Finally, characteristics of the communities in which young women live may play a role in shaping the kinds of opportunities they have to spend time in specific activities. The HPTN 068 study population is enrolled within a well-defined research site across 24 villages. We are specifically interested in exploring whether the alcohol outlet density of each village influences sexual risk as measured by individual-level prevalent herpes simplex virus 2 (HSV-2) infection of village residents. Much previous research has been focused on identifying individual sexual risk behaviors, but the structural and place-based characterization of sexual risk has been largely understudied, particularly within our target population. The results of these analyses may provide insight into distal determinants of sexual risk outcomes as well as produce information that could be used to better inform and target preventive interventions.

A.2 Specific Aims

Aim 1: Estimate the effect of **school enrollment** on **pregnancy rates** among rural South African young women.

Hypothesis: We hypothesize that pregnancy rates will be lower among those enrolled in school compared to those not enrolled in school.

Methods: We will construct a cohort from the Agincourt Health and Demographic Surveillance Site of nulliparous young women aged 12 to 18 followed between 2000 and 2012. In this longitudinal analysis, we will build Cox proportional hazards models to estimate the hazard ratio comparing the pregnancy rates among young women currently enrolled in school compared to those not currently enrolled in school. We will treat this school enrollment exposure as time-varying and will control for important time-varying and time-fixed confounders.

Aim 2: Estimate the association between **frequency of visits to alcohol outlets** and sexual risk behavior (**sexual debut and unprotected vaginal sex acts**) among rural South African young women.

Hypothesis: We hypothesize that increasing number of alcohol outlet visits will be associated with increased likelihood of sexual debut and increased number of unprotected sex acts.

Methods: In this cross-sectional analysis, we will build a logistic regression model to estimate the association between number of alcohol outlet visits in the last six months and sexual debut. We will

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

build a zero-inflated Poisson regression model to estimate the association between number of tavern visits and number of unprotected vaginal sex acts in the last three months. Both exposure and outcome variables are self-reported in the baseline questionnaire for all young women enrolled in HPTN 068.

Aim 3: Estimate the association between **alcohol outlet density** and **prevalent HSV-2 infection** among rural South African young women.

Hypothesis: We hypothesize that living in villages with higher alcohol outlet density will be associated with higher likelihood of prevalent HSV-2 infection.

Methods: Using a multilevel logistic regression model to control for individual- and village-level risk factors, we will estimate the association between village-level alcohol outlet density and individual HSV-2 sero-status. HSV-2 sero-status will be measured at baseline for each young woman enrolled in HPTN 068.

B. BACKGROUND AND SIGNIFICANCE

Engaging in risky sexual behaviors can lead to numerous negative health outcomes including unintended pregnancy, HIV infection, and infection with other STIs. A better understanding of the determinants of sexual risk is critical to informing new preventive interventions to reduce the burden of these health outcomes. We hypothesize that an individual's routine activity pattern can influence her sexual risk. Below, we justify this hypothesis by: briefly summarizing the current public health burden of sexual risk outcomes with a focus on young women in South Africa; setting up the conceptual and theoretical framework for the association between routine activity patterns and sexual risk; and discussing the literature to-date on the association between school enrollment, visits to alcohol outlets, alcohol outlet density, and sexual risk.

Overview

The public health burden of HIV remains high throughout the world, particularly among young South African women. There are currently 34.0 million people living with HIV worldwide, according to most recent estimates, and the number has risen continuously since 1990.(1) South Africans face a burden from HIV/AIDS at a level largely unmatched in the rest of the world. Overall, the national HIV prevalence was estimated to be over 15% in 2011 with over 5.5 million people living with the disease.(2) Young women in South Africa are at exceptionally high risk for HIV infection. They are infected at a rate over three times that of young men, and, by age 21, nearly a third of young women are HIV positive.(3, 4) The question remains: Why are these young women so strongly affected by HIV and what can be done to prevent the spread of the disease in this vulnerable population?

Similarly, there is a significant public health burden associated with teenage pregnancy; this burden is particularly high among adolescents in sub-Saharan Africa. Teenage childbearing has been associated with negative health and social outcomes for both teenage mothers (maternal mortality, lower self-esteem, lower educational attainment, and higher poverty levels) and their offspring (increased risk of preterm delivery, low birth weight, and infant mortality).(5-7) The negative consequences for children of teen mothers persist throughout their own childhood and early adulthood as they also tend to have higher poverty levels, attain lower educational levels, and are at higher risk of becoming teen parents themselves.(5) Globally, over 16 million births occur to mothers under the age of 20 every year, and nearly all (95%) occur in developing countries.(8) In South Africa, over 10% of all births are to teen mothers(9), and, by the time they reach age 20, nearly one-third of all young women will have begun childbearing.(10) Successful preventive interventions for teenage pregnancy, particularly in areas where it is most common, could potentially prevent a large burden of negative health outcomes.

Much previous research has been conducted to identify individual-level characteristics associated with sexual risk outcomes. In particular, increased risk of HIV has been associated with demographic factors such as age (11), gender (12), and socioeconomic status (13); and behavioral factors such as early coital debut (14), number of sexual partners (15), condom use (15), and other STI diagnosis (16). Becoming a teenage mother has been associated with demographic factors such as race/ethnicity, large household size, and not living with biological father; and behavioral factors such as frequent sex and low contraceptive use.(7, 17, 18) However, many of these individual-level characteristics are difficult to modify, or completely non-modifiable, with little use in informing preventive interventions. Although more distal risk factors such as poverty and low education have also been associated with both HIV (13, 19, 20) and teenage pregnancy (21, 22), other structural and contextual risk factors, particularly those with the potential to inform interventions, have been explored much less frequently.

This study aims to explore how routine activity patterns (school enrollment and frequency of visits to alcohol-serving establishments) and community characteristics (alcohol outlet density) may influence sexual risk. We draw a distinction between activities that provide structure and supervision to adolescents (school enrollment) and activities that allow for periods of unstructured and unsupervised time (visits to alcohol-serving establishments). In general, we hypothesize that structured and supervised activities will be associated with lower sexual risk and unstructured and unsupervised activities will be associated with higher sexual risk.

Conceptual and theoretical framework

Figure 1 is a conceptual diagram of the association between routine activities and sexual risk, representing the contextual factors that may influence routine activities and the potential mechanisms through which we hypothesize routine activities may influence sexual risk. The first column depicts the contextual factors we hypothesize influence activity participation. The propensity for an individual to participate in any specific activity is likely shaped by factors at the community-level (does the community provide opportunities to participate in the activity?), household-level (does the family encourage or discourage certain activities, either directly or because of characteristics associated with participation in the activity?), and individual-level (does the individual have certain characteristics that are associated with participation in the activity?).

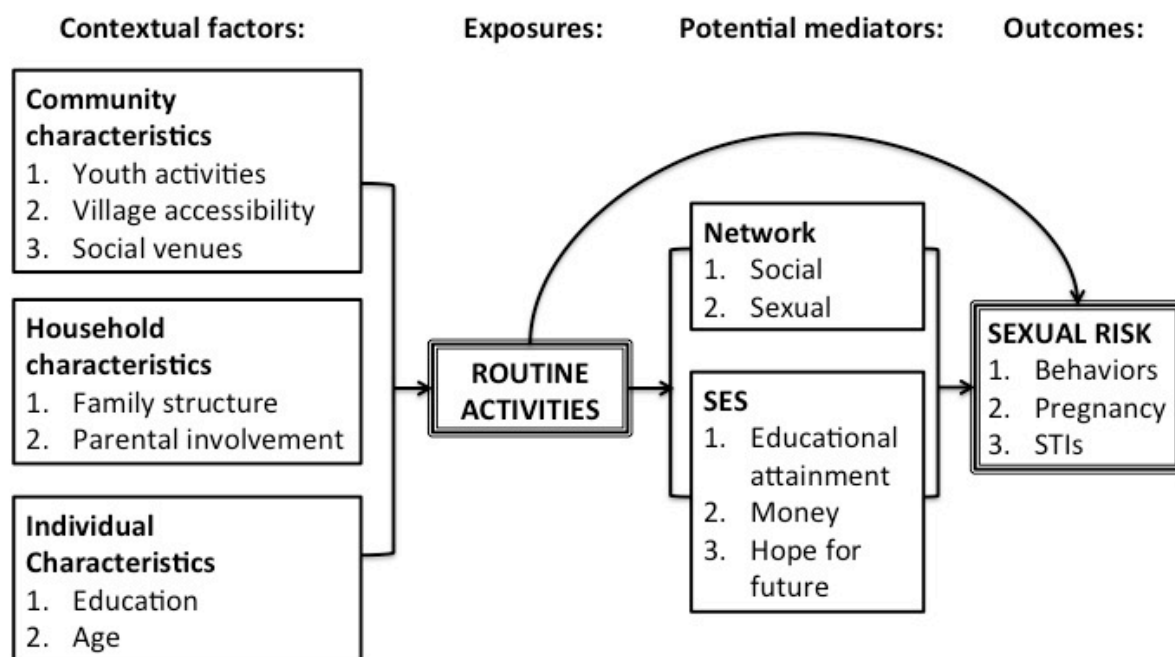


Figure 1. Conceptual diagram of the influence of routine activities on sexual risk

The third column depicts our hypotheses that routine activities can influence sexual risk through three separate, though not mutually exclusive mechanisms. First, **routine activities may influence sexual risk outcomes directly by providing periods of structured and supervised time during which sexual risk behavior is unlikely to occur, or by providing periods of unstructured and unsupervised time during which sexual risk behavior may be more likely to occur.** The justification for this proposed direct effect is grounded in a synthesis of Social Control Theory (23, 24), and the routine activity/time-use perspective (25-28). Both of these theories have been used primarily in criminology to explain

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

delinquent behaviors, but their applicability can be extended to sexual behavior.(29) In general, Social Control Theory posits that individuals may choose to exhibit a wide range of behaviors, but this range can be limited by social forces. The routine activity perspective adds to this framework by proposing that activities involving unstructured socializing with peers (in the absence of authority figures) are the least limited by social forces, and allow time and opportunity for deviant behavior to occur.

Second, participation in specific activities may change one's social and sexual network. If the co-participants are generally safer (ex: younger, more educated) and become integrated into the network that influences a participant's sexual behavior, then one might expect decreased sexual risk outcomes for the participant. If the co-participants are generally riskier (ex: older, less educated), then one might expect increased sexual risk outcomes for the participant. The justification for this proposed mechanism is grounded in Social Cognitive Theory.(30) This theory posits that people learn information and change their own behaviors by viewing others. Peers can be strongly influential providers of sexual information and sexual norms to adolescents; this is particularly true given that more traditional sources of general information (parents and schools) often do an inadequate job of frankly discussing sexual matters with adolescents.(31) These peer groups, thus, can be an important source of social support for positive behavior modeling or negative behavior modeling, depending on the norms of the group.

Finally, participation in specific activities may improve one's socio-economic status (SES), which could offset some of the financial motivations to engage in risky sexual activity. This mechanism may not be applicable to all routine activities, but certain activities, such as school enrollment, as well as other skill-building and educational activities, may improve the participant's future job prospects and hopes for the future. However, the timeframe in which we might anticipate this kind of mechanism to influence sexual risk is likely to be long in duration. Since we are studying more temporally-proximate outcomes in a population of young women, this mechanism is not explored further here.

Here, we will briefly link our specific research aims to this conceptual framework; a literature review for each aim will be presented in subsequent sections below. In both Aims 1 and 2, we are interested in estimating the association between individual routine activities and sexual risk outcomes. In Aim 1, we are investigating the association between the routine activity of school enrollment and the sexual risk outcome of teenage pregnancy. In Aim 2, we are investigating the association between frequency of visits to alcohol outlets and sexual risk behavior. The conception and design of both of these aims is grounded in the theories listed above used to explain the potential mechanisms of the association between routine activities and sexual risk, as presented in Table 1.

In Aim 3, we are investigating the association between community-level alcohol outlet density and the individual-level sexual risk outcome of HSV-2 sero-prevalence. The design and analysis of this aim will be grounded in a Neighborhoods and Health theoretical framework.(32) This framework was developed to explore how neighborhood factors influence the health of residents, with emphasis on the idea that there exist neighborhood-level contextual factors (characteristics of the neighborhood itself, such as service availability or green spaces) that can influence individual health above and beyond the neighborhood-level compositional factors (aggregate individual-level factors such as median income and unemployment rate).(33)

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

Table 1. Theoretical support and exposure measures by study aim

Aim	Theoretical support	Exposure measure
1. School enrollment and teen pregnancy	1. Social cognitive theory 2. Social control theory	School enrollment status
2. Visits to alcohol outlets and sexual behavior	1. Social cognitive theory 2. Social control theory	Number of visits to taverns in last six months
3. Alcohol outlet density and HSV-2	Neighborhoods and Health	Village alcohol outlet density

Aim 1: School enrollment and teen pregnancy

An important structured place where many young women may spend time is school. In South Africa, nearly all young people are enrolled in school for most of their school-aged years; by the time they reach secondary school age, the vast majority (94%) still remain in school.(34) However, school drop-out rates begin to rise after the mandatory enrollment age of 16.(35) The average school year in South Africa has 200 school days and students typically spend five to six hours a day at school.(36, 37) It is important to note, however, that the quality of education in South Africa is substandard with reports of low performance in literacy and mathematics, high failure rates, high drop-out rates, and underqualified teachers, particularly in poor, majority black schools.(38, 39) Nonetheless, schooling may provide large periods of structure and supervision in young adults' lives, as well as teach important life skills, provide sex education, and increase future opportunities.

Educational attainment is associated with sexual risk. Lower educational attainment has been associated with sexual risk behaviors such as early sexual debut, unprotected sex, and multiple partnerships throughout sub-Saharan Africa (40-42). Lower educational attainment has also been linked to teenage childbearing.(40) Finally, lower educational attainment has been associated with HIV incidence (43) and HIV prevalence (19, 41, 44). However, there is evidence that this relationship has changed over time, as, earlier in the epidemic, higher HIV risk was associated with higher education levels.(45-49)

Studies of educational attainment and sexual risk are compromised by incomplete exposure definition and weak study designs. The highest grade reached by an individual in her lifetime may provide a good measure of the education to which she has been exposed, the life skills she has learned, and/or the increased socio-economic status that comes from increased education.(50) However, measures of educational attainment do not incorporate the potential protective effects of structured time or safer networks that may result from school enrollment. This is particularly true in studies of relatively older populations who may be long out of school.(43) Further limiting the findings from the current research is that most of these studies were performed cross-sectionally (19, 40, 41, 44) so inferences about the directionality of the relationship between education and sexual risk are limited.

School enrollment status is also associated with sexual risk. An alternative education measure to overall educational attainment is school enrollment status. School dropout has been linked to sexual risk behaviors like multiple partnerships, partner age, unprotected sex, and transactional sex.(51, 52) School enrollment status is also associated with lower HIV prevalence, though no studies to date have linked enrollment status to HIV incidence.(51, 53) The relationship between school enrollment and teenage pregnancy is more complex. There is evidence of an association between pregnancy and subsequent school drop-out.(54-56) However, the reverse association may also be true. Results from two interventions designed to incentivize young women in Sub-saharan Africa to stay in school longer report

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

lower teenage pregnancy rates(57, 58), supporting similar results from a single longitudinal study in the United States from 1988-1994.(59)

Due to several methodological limitations, the nature of the relationship between school enrollment and sexual risk remains unclear. First, as with educational attainment, the majority of studies linking school enrollment and sexual risk are cross-sectional in design (51-55), limiting the ability to make inferences about the directionality of the observed associations. However, two of these studies did try to address this issue by using detailed calendars to record event histories.(54, 55) A more appropriate way to establish the directionality of the association between school enrollment and sexual risk would be to analyze data prospectively collected in a longitudinal dataset. Relatedly, in order to establish the correct sequence of events, the timing of changes in school enrollment status and timing of sexual risk outcomes need to be recorded finely. This is often difficult in cross-sectional studies relying on self-reported data, particularly those asking respondents to remember temporally-distant events. More objective measures of exposures and outcomes (such as school enrollment and delivery dates recorded in a yearly census) could provide better information to understand this complex relationship. Responding to the methodological limitations and general sparseness of the current literature, Jukes, *et al* state that there is an “urgent need to strengthen the evidence base” on education and sexual risk.(50) Using a large, longitudinal, census-based dataset, this study aims to address many of the above-mentioned limitations and add to the small, yet growing body of literature on the association between school enrollment and sexual risk.

Aim 2: Frequency of visits to alcohol outlets and sexual risk behavior

Frequenting alcohol outlets may influence sexual risk behaviors. Visits to alcohol outlets (establishments where alcohol is sold) may allow for periods of unsupervised and unstructured time which we hypothesize is associated with increased sexual risk behaviors. Additionally, the alcohol consumption associated with visits to social venues like bars and taverns could directly influence sexual risk behaviors. We discuss both of these potential mechanisms below, but begin with the significance of using sexual behaviors as outcomes.

Sexual behaviors are important proximate determinants of sexual risk outcomes. Although we typically think of sexual risk outcomes as disease or unintended pregnancy, the behaviors that put young women at risk for these events can also be important to explore. In particular, becoming sexually active and having unprotected sex have a clear relationship with HIV, other STIs, and pregnancy as together they are the mechanism for sexual transmission of disease and conception. Unprotected sex has been strongly linked with STI infection (60-63) and pregnancy rates.(64) Young women in South Africa do not report particularly risky sexual behaviors relative to their counterparts in the United States, but in the context of heavy HIV/STI burden, any sexual activity can be risky.(65) Identifying factors associated with these sexual behaviors is important to better understand the etiology of sexual risk and to develop potential prevention interventions.

Alcohol use and abuse is associated with sexual risk. The link between alcohol use and sexual risk is well-established in populations throughout the world (66-70). This association has been observed regionally among populations in Sub-saharan Africa (71-73) and specifically in the Republic of South Africa.(74, 75) These findings appear to hold across both genders, over a range of ages, for both behavioral outcomes such as unprotected sex(66, 68, 71), and biologic outcomes such as HIV infection.(69, 70, 72) Additionally, these associations appear to be robust to different measures of alcohol use, holding among global exposure studies (‘What is the frequency and quantity of recent

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

alcohol consumption?’), event-level exposure studies (‘Was the individual under the influence of alcohol during particular occasions of risky sex?’), and studies where exposure was randomized.(68, 76)

Biologic evidence supports the claim for a causal relationship between alcohol use and sexual risk.

Consumption of alcohol leads to disinhibition and impaired decision-making (73, 77) and decreases awareness of social norms and risk-inhibiting cues (78). Relatedly, alcohol use leads to a myopic focus on the present, where distant consequences of proximal activities are discounted.(78, 79) These disruptions to the normal social and individual constraints on risky activity are thought to lead to increased sexual risk. This negative effect is reinforced by the fact that alcohol users themselves report feelings of enhanced sexual appeal and reduced sexual control.(75) Finally, there is also evidence for a dose-response relationship between alcohol use and sexual risk as, generally, greater sexual risk is associated with increased quantity of alcohol consumption.(71)

Typical drinking patterns in South Africa are dangerous. Approximately 25% of men and 10% of women experience alcohol problem symptoms in South Africa.(80) There is also evidence for high levels of alcohol misuse among South African young adults; nearly a quarter of all high school students report recent binge drinking activity.(80, 81) Within the WHO-based regional grouping for southern Africa of which South Africa is a member, an estimated 55% of males and 33% of females drink alcohol. For each drinker in the region an estimated 16.6 liters of alcohol is consumed per year, placing it among the few areas in the world with the most detrimental drinking patterns.(82)

Alcohol use and misuse are associated with patronizing alcohol outlets. In South Africa, Morejele *et al.* have developed a conceptual model of the relationship between alcohol use and sexual risk behaviors informed by data collected using qualitative methodology.(83) They found that the environment in which drinking occurs is a moderating factor in the association between alcohol use and sexual risk. This proposed relationship is supported by quantitative data linking patronizing alcohol outlets to increased frequency and quantity of drinking in South Africa.(84)

Aside from the biologic effects of alcohol consumption, the context of the places where alcohol is served may also influence sexual risk. The physical features of a drinking establishment may be conducive to unprotected sex occurring on premises. Qualitative evidence suggests that the music, dim lights, uni-sex toilets, and lack of condoms typical in bars create favorable environments for risky sexual activity.(83) Additionally, there is evidence that the network of people to whom an individual is exposed in drinking establishments may be riskier than in non-drinking environments. Sexual partners are often met in these places, particularly for young women, and these partners are often older men looking for sex in exchange for money.(73)

Risky social venues that sell alcohol have been identified using a community mapping methodology.

Specific social venues have been identified as high-risk using the Priorities for Local AIDS Control Efforts (PLACE) methodology throughout the world.(85-94) This methodology identifies venues where people meet new partners using qualitative interviews and mapping exercises with key informants.(95) In South Africa, more than 85% of the places identified in this manner were alcohol-serving establishments.(85) These high-risk locations are appealing targets for interventions; however, interventions directed at PLACE-identified locations have had mixed results.(93, 94)

Direct evidence for the association between visits to alcohol outlets and sexual risk is sparse. Although the sexual risk profiles of participants recruited at alcohol outlets have been relatively well-characterized as risky(73, 96-99), studies comparing the sexual risk profiles of those with different levels

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

of alcohol outlet exposure are much fewer. Two cross-sectional studies in Zimbabwe report moderate associations between recent beerhall visits and prevalent HIV infection among men(100) and prevalent HIV infection and risky sexual behavior among both men and women.(101) A single cross-sectional study of an adult population in South Africa found that patronizing an alcohol outlet was associated with increased unprotected vaginal sex, and an association with increased sexual risk index was significant even after controlling for alcohol consumption.(84) A final, cross-sectional study in India found an association between patronizing wine shops and prevalent STI infection (HIV, HSV-2, syphilis, chlamydia, or gonorrhea) among adult males.(102) However, interventions targeted in these high-risk venues have so far had weak or null results on sexual outcomes.(103, 104) Studies linking visits to specific locations to sexual risk in adolescents are rare, though one novel study used risk mapping and focus groups among secondary students in Zimbabwe to identify six high-risk area categories, including commercial centers where alcohol is served.(105)

The current literature on exposure to alcohol outlets and sexual risk is limited, particularly for young women in sub-Saharan Africa. Only four studies have empirically examined this association and, although they all four find significant results in adult populations, this association has not been confirmed in youth populations in general, or among South African young women in particular. Also, the South African and Indian studies used convenience sampling methods where many members of the study population were enrolled within alcohol outlets themselves which could cause bias if they were not representative of the general population(84, 102). Finally, all four studies dichotomized the drinking establishment exposure into patrons versus non-patrons; this dichotomy may mask interesting nuances in the relationship between amount of exposure to drinking establishments and sexual risk. Clearly, more research is needed to add to the small body of literature linking visits to drinking establishments and sexual risk.

Aim 3: Alcohol outlet density and sexual risk

Characteristics of the communities in which young women live may play a large role in shaping their exposure to different risk environments. Communities that provide safe resources and opportunities for youth may directly or indirectly lower their sexual risk while communities that provide opportunities for youth to engage in unsupervised activities may directly or indirectly increase their sexual risk. We hypothesize that alcohol outlets provide opportunities for youth to engage in unsupervised activities and that living in communities with more alcohol outlets will be associated with increased sexual risk.

Specific community characteristics are associated with sexual risk, but modifiable community characteristics are understudied. Associations with increased sexual risk have been noted for multiple sociodemographic, behavioral and spatial factors, including: urban location and age/gender profiles(106); urban status and ethnicity/unemployment profiles(107); proximity to a city and HIV prevalence within a 25 kilometer buffer(108); GDP, mobility, urbanization and drug use(109); proximity to a primary road(110, 111); proximity to a market and distance to a health clinic(112); and literacy rates, unemployment, poverty, and urban residency.(113) However, most of the community-level exposures explored to-date are non-modifiable or difficult to modify. Characterizing neighborhoods based on their resources and venues, like alcohol outlets, would be helpful in terms of prevention potential, but has been largely ignored in research seeking to identify determinants of sexual risk.

Alcohol outlet density is associated with alcohol consumption and alcohol-related problems. The association between increased alcohol outlet density and increased frequency and consumption of alcohol consumption is well-established.(114-117) Increased alcohol outlet density is also associated with a variety of alcohol-related problems, including automobile injuries(118) and violent crime.(119)

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

Public policy and economic frameworks also support the link between alcohol outlet density, alcohol consumption, and alcohol-related problems. These perspectives posit that people tend to consume less alcohol and therefore have fewer alcohol-related problems when the accessibility to alcohol is limited, either through higher legal purchasing age, increased prices, or decreased outlets.(120-122) It should be noted, however, that the research to date on alcohol outlet density is heavily skewed to populations in the developed world. The vast majority of studies were performed in American study populations and, to our knowledge, the exposure of alcohol outlet density has been ignored in South African study populations.

Community alcohol outlet density may be associated with sexual risk as well. To our knowledge, there have been four studies examining the association between alcohol outlet density and sexual risk(123-126). Increasing alcohol outlet density was significantly associated with census-level gonorrhea rates in a cross-sectional study in Louisiana(124) and a time series analysis in California(123), and with self-reported STI infections in a cross-sectional study in Louisiana and California.(125) A final cross-sectional study in Namibia found that increasing alcohol outlet density was associated with increased HIV prevalence.(126) This evidence suggests that the alcohol environment in a neighborhood can influence risk. However, in South Africa, researchers found that proximity of an alcohol outlet to home residence was inversely associated with increased sexual risk behavior among female patrons of alcohol outlets, indicating that the anonymity that comes from visiting more distant venues may have a disinhibiting effect.(127)

Current research on the association between alcohol outlet density and sexual risk is methodologically limited. First, studies examining both exposures and outcomes at the community-level may be subject to the ecologic fallacy if inferences are made at the individual-level.(123, 124, 126) However, if studies, instead, examine the influence of community-level factors on individual-level outcomes, appropriate statistical methods like multilevel modeling or generalized estimating equations must be used to account for the fact that observations in the same community are likely non-independent.(125) Finally, the cross-sectional nature of nearly all of the studies also limits the ability to assess directionality of the association between exposure and outcome.

The evidence to-date on alcohol outlet density and sexual risk is incomplete. All but one of the four previous studies were conducted in the United States and all four were conducted in adult or general study populations. Replication of these results in youth-specific populations outside of the United States, such as the high-risk population of South African young women, would further extend the generalizability of the results. This is particularly pertinent as youth may have different alcohol outlet utilization patterns and alcohol-using behaviors than adults, so generalization from adult population studies to youth could be problematic.

Conclusion

This proposed study has important public health implications. Identifying activity patterns associated with high and low sexual risk will help us further understand the etiology of sexual risk outcomes like HIV and teen pregnancy and potentially provide insight into how sexual behavior (both safe and unsafe) is learned and transmitted. Also, the results of this study could directly inform prevention interventions. If particular activities are found to be protective against or predictive of sexual risk outcomes, interventions to increase or decrease access to those activities could be developed and implemented relatively easily. Using alcohol outlets as targets for prevention interventions is particularly attractive.(85, 128) The development of better prevention interventions for South African young women

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

will be critical to prevent the continued spread of disease and high burden of teenage pregnancy in this vulnerable population.

C. RESEARCH DESIGN AND METHODS

C.1 Study Overview

Overall, the objective of this research project is to explore the association between routine activities on sexual risk among young women in rural South Africa. We will approach this research question in three ways. In the first aim we will use longitudinal data routinely collected from the Agincourt Health and Demographic Surveillance Site (HDSS) to explore the influence of school enrollment on teenage pregnancy. The second aim will be performed as an ancillary study to HPTN 068 - The Effects of Cash Transfer and Community Mobilization in Young South African Women Study. Here, we will explore the association between frequency of visits to alcohol outlets and self-reported sexual risk behavior. The third aim will explore the association between village-level alcohol outlet density on sexual risk using geographic information collected by the Agincourt HDSS, and village resource and individual biologic data collected by HPTN 068.

C.2 Study population and Parent studies

To conduct the first aim of this study, we will use previously collected data from the rural Bushbuckridge Sub District in the Mpumalanga province of South Africa. The Medical Research Council / Wits University Rural Public Health and Health Transitions Research Unit runs the Agincourt HDSS in this area and a complete census is performed on all the households yearly. The HDSS is a multi-round, prospective, community study established in 1992; since then it has continually monitored all vital events occurring within the defined geographic area. The HDSS now covers 90,000 people living in 16,000 households in 24 villages.(129) Cohort profiles and details of the data collection methods have been described previously.(129, 130) Figure 2 displays a map of the study site location and the spatial distribution of study villages.

The second and third aims will use baseline data originally collected by HPTN 068. HPTN 068 is a Phase III randomized trial currently being conducted within the Agincourt HDSS and surrounding area. The study aims to determine whether providing cash transfers conditional on school enrollment reduces HIV risk in young women. Eligibility inclusion criteria at baseline are: female; age 13-20; enrolled in grades 8, 9, 10, or 11 at a school in the HDSS study site; intention to live in the study site until the end of follow-up; consent/assent to HIV and HSV-2 testing; ability to read sufficiently to self-administer a computer assisted interview; have documentation to be able to open a bank account; and live with a parent/guardian willing to consent to all study procedures. Eligibility exclusion criteria are: pregnancy at baseline; married at baseline; and no parent or legal guardian in the household. Importantly, young women who test positive for HIV at baseline will not be excluded from study participation. HPTN 068 is currently in its second year of follow-up and has a total enrolled study population of 2536 young women.

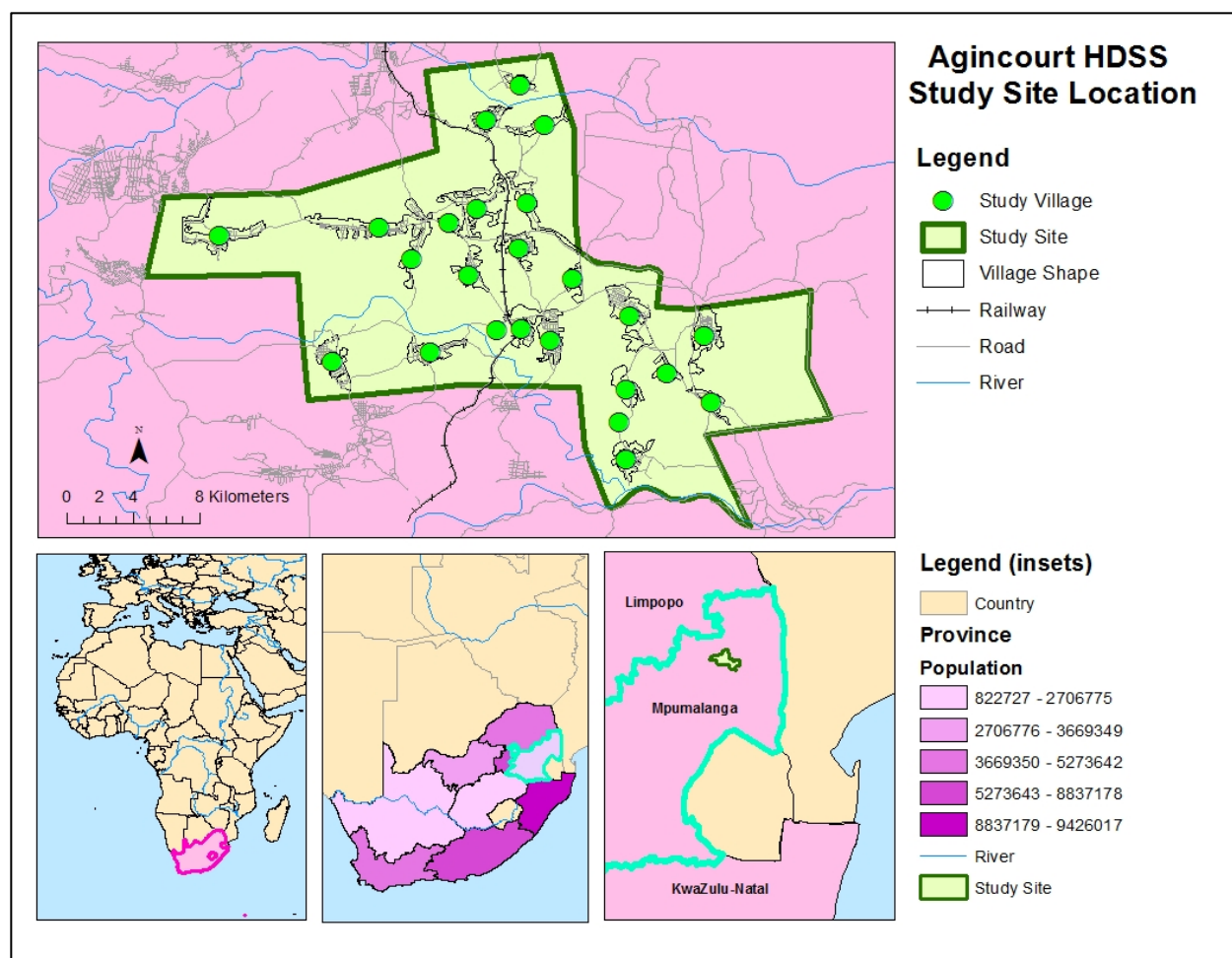


Figure 2. Map of the study site location

C.3 Aim 1: School enrollment and teenage pregnancy

C.3.1. Study design

To investigate the association between school enrollment and pregnancy, we propose a study that is longitudinal in design with both the exposure and the outcome assessed and analyzed at the individual-level. We will construct a cohort from data originally collected by the HDSS that will include up to six years of follow-up for each participant. School enrollment exposure will be included as a time-varying exposure and each participant will be followed up retrospectively for incident first pregnancy.

C.3.2 Data collection and variable measurement

Cohort construction: We will construct the cohort using data from the HDSS census. All 12-18 year-olds who had not yet experienced their first pregnancy and who lived in the study area between 2000-2012 contributed person-time to this study. Nulliparous young women who moved into the study site after their 12th birthday and nulliparous young women older than 12 in 2000 will be treated as late entries. Young women who move out of the study site or die during follow-up will be censored at the time of move or death. All young women who do not experience their first pregnancy by the age of 18 will be administratively censored on their 18th birthday.

Outcome assessment: The outcome we are exploring in Aim 1 is **time to first pregnancy**. All vital events that occur in the HDSS study area, including births, deaths, and migrations, have been updated

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

continuously in the database since 1992. In the yearly census update, fieldworkers administer a pregnancy module to collect information on all pregnancies that took place in the previous 12 months. Although this list of pregnancies includes those that did not result in live births, this outcome is self-reported, and miscarriages and abortions, particularly early losses, may be under-reported. We anticipate that unreported early pregnancy losses in this population will be minimal and the unreported losses will be non-differential by school-enrollment status; these assumptions will be tested with sensitivity analyses.

Date of delivery and estimated length of pregnancy are recorded for all pregnancies. We will compute the estimated conception date for each pregnancy reported by subtracting the self-reported length of the pregnancy from the date of delivery. The time to first pregnancy will be calculated as the duration, in months, between a young woman's 12th birthday and the estimated conception date of her first pregnancy. Self-reported duration of pregnancy could be subject to recall bias, which would be problematic if incorrect recall was associated with school enrollment status. However, given that pregnancy duration has natural bounds that are relatively short (approximately between a minimum of 4 and a maximum of 42 weeks) and our overall time scale is very long (12 years), we do not expect slight over- and under-estimates of pregnancy duration to notably bias our results. All pregnancy duration data will be cleaned and monitored for outliers outside of the natural bounds of possible duration.

We restricted our outcome to first pregnancy as opposed to pregnancy of any order for several reasons. First, given the relatively young ages in our cohort, we do not expect to see many young women who will go on to have multiple pregnancies during follow-up. Second, we hypothesize that young women who have previously had a child will become incomparably different from those who have not, particularly with regards to their school enrollment. Specifically, young women who have had a child are less likely to go on and have further school enrollment exposure, even though previous pregnancy is not completely incompatible with future schooling, particularly in South Africa.(56)

Exposure assessment: The primary exposure we are exploring in Aim 1 is **school enrollment**. School enrollment will be coded as a time-varying bivariate exposure that equals one when the participant is enrolled in school and zero when not enrolled in school. An education status module was administered in 1997, 2002, 2006, 2009, and 2012. At the time of each module, field workers recorded the highest education level each young woman had achieved and whether or not she was currently enrolled in school. Although these modules were only updated, on average, every three years, we can interpolate school enrollment at any given time by examining change in educational attainment between module years. For example, if a young woman reports current enrollment in Grade 8 in 2002 and current enrollment in Grade 12 in 2006, we can infer that she has been continuously enrolled in school between 2002 and 2006. Table 2 shows other potentially more complex data configurations and the *a priori* coding decisions we will make about school enrollment duration. We will perform sensitivity analyses around these assumptions to evaluate the robustness of our results to changes in school enrollment exposure coding. To confirm the exposure status of those who go on to have a pregnancy during follow-up, we will also use information from the pregnancy module that includes a question asking if the mother was a student at the time she became pregnant.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

Table 2. Example data configuration possibilities with school enrollment exposure coding decisions

School enrollment data configuration possibilities					Exposure duration	Notes
1997	2002	2006	2009	2012	(2000-2012)	
Not enrolled None	Not enrolled None	Not enrolled None	Not enrolled None	Not enrolled None	0 years	
Enrolled Grade 5	Enrolled Grade 10	Not enrolled Grade 12	Not enrolled Grade 12	Not enrolled Grade 12	4 years (2000-2004)	
Enrolled Grade 5	Enrolled Grade 10	Not enrolled Grade 10	Not enrolled Grade 10	Not enrolled Grade 10	2 years (2000-2002)	
Missing	Enrolled Grade 8	Missing	Not enrolled Grade 10	Not enrolled Grade 10	4 years (2000-2004)	Infer over missing data in middle
Missing	Missing	Enrolled Grade 10	Not enrolled Grade 10	Not enrolled Grade 10	6 years (2000-2006)	Infer over missing data in beginning
Enrolled Grade 3	Enrolled Grade 8	Missing	Missing	Missing	2+ years (2000-?)	MI to impute educational attainment
Enrolled Grade 3	Enrolled Grade 8	LTFU (2005)	LTFU	LTFU	4 years (2000-2004)	Enrollment status extended to time of LTFU
Enrolled Grade 3	Not enrolled Grade 7	LTFU (2005)	LTFU	LTFU	1 year (2000-2001)	Enrollment status extended to time of LTFU
Enrolled Grade 3	Enrolled Grade 8	Enrolled Grade 10	Not enrolled Grade 10	Not enrolled Grade 10	6 years (2000-2006)	Assume continuous enrollment, but flag for temporary dropout or grade repetition

Covariate assessment: We will be exploring the potential influence of several key covariates on the relationship between pregnancy and school enrollment. Specifically, we will request data on:

- 1) **Age**, calculated, in years, from birthdate. We hypothesize that increasing age will be associated with an increased likelihood of pregnancy and a decreased likelihood of school enrollment.
- 2) **Educational attainment**, defined as a dichotomous, time-varying covariate that equals one if the participant has graduated from secondary school, and zero otherwise. We hypothesize that educational attainment acts as a mediator between school enrollment and pregnancy.
- 3) **Village**, measured as the village of residence at time of study entry. We hypothesize that certain villages may have different school enrollment environments. We also hypothesize that village may be associated with pregnancy indirectly through its association with household SES.
- 4) **Gender of household head**, defined as the gender (male or female) of the individual reported as the head of household. We hypothesize that living in a household headed by a female will be associated with decreased likelihood of school enrollment (through decreased parental monitoring and household SES) and increased likelihood of pregnancy (through decreased household SES).
- 5) **Household size**, defined as the total number of individuals reporting membership in the participant's household at the time of study entry. We hypothesize that increasing household size will be associated with decreased school enrollment (through decreased parental monitoring and household SES) and increased likelihood of pregnancy (through decreased household SES).
- 6) **Household socioeconomic status (SES)**, measured as a time-varying composite index of general SES based on household assets (SES module administered in 2001, 2003, 2005, and 2007). We hypothesize that increasing household SES will be associated with an increased likelihood of school enrollment and decreased likelihood of pregnancy.
- 7) We will also explore the influence of **calendar year** on the relationship between school enrollment and pregnancy by creating an indicator for the year in which each participant turns

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

12 years old. As this is a long follow-up period, it is possible that both teen pregnancy rates and school enrollment rates vary over time due to external factors.

The hypothesized relationships between the covariates, school enrollment and pregnancy are diagrammed in Figure 3.

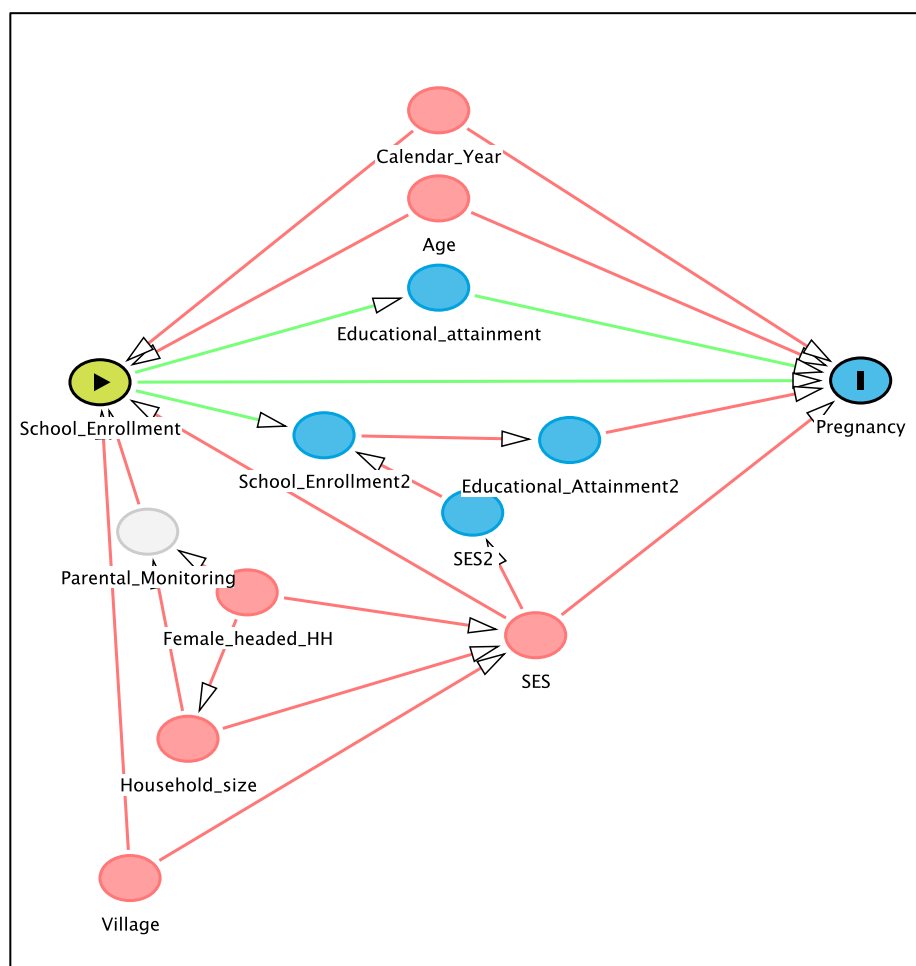


Figure 3. Directed acyclic graph depicting the hypothesized relationship between school enrollment and pregnancy

C.3.3 Statistical approach

Using a Cox proportional hazards model, we will compare the hazard of first pregnancy among those enrolled in school compared to those not enrolled in school. The origin for each participant will begin on her 12th birthday and the time scale will be age. We will partition the dataset so that young women who switch school enrollment status during follow-up can contribute both exposed and unexposed person-time. If a young woman's estimated conception date falls mid-way through the final school year she reports attending, we will confirm enrollment status at pregnancy by deferring to the validation question: "Were you a student at the time you became pregnant?". Young women will be censored at the time of their first pregnancy, if they move, die, or are otherwise lost to follow-up, or when they reach 18 years of age.

We will assess for the potential confounding effects of time-fixed (baseline village of residence, gender of household head, household size, year of study entry) and time-varying (household SES) covariates.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

Age will be adjusted for implicitly as age is the time scale in this analysis. We will identify a minimally sufficient adjustment set using the directed acyclic graph (Figure 3). Only covariates in the minimally sufficient adjustment set that alter the point estimate by greater than 10% will be retained in the final adjusted model. We will also consider the potential for effect measure modification by calendar year and age. To assess for effect measure modification, we will categorize the potential modifier and examine the stratified estimates visually to assess whether there are differences in the magnitude of effect across strata. To formally test for effect measure modification, we will include an interaction term between the potential modifier and the school enrollment exposure variable in the model. Interaction terms with Wald test statistic p-values of <0.10 will be considered significant and stratified estimates will be reported.

C.2.4 Power

We parameterized our power calculation using the publically available ten percent sample of the full HDSS dataset. In this sample, a total of 2097 young women between the ages of 12 to 18 contributed 6308 person-years between 2000 and 2012. This number of person-years incorporates loss to follow up due to death or move outside of the study area, aging in and out of the cohort, and censoring after first pregnancy. We therefore assume 63,080 person-years will be observed in the full cohort. Approximately 50% of the sample reported being enrolled in school at each of the time points; we assume that 50% of the person years under study in the full sample will be exposed to school enrollment as well. A total of 152 first pregnancies (7.3%) were recorded in the ten percent sample; we assume we will observe approximately 1520 first pregnancies in the full sample.

Using an expected data approach for a simple Poisson model with a two-sided test with an alpha of 0.05, we calculated the 95% confidence limits across the range of potential protective hazard ratios from 0.50 to 0.90 (Table 3). All expected confidence intervals exclude the null and are precise with confidence limit ratios around 1.25. We maintain a statistical power greater than 80% for hazard ratios from 0.50 to 0.86, as confirmed in a power calculation using the same parameters estimating the expected power for all hazard ratios between 0.7 and 0.9 (Figure 4).

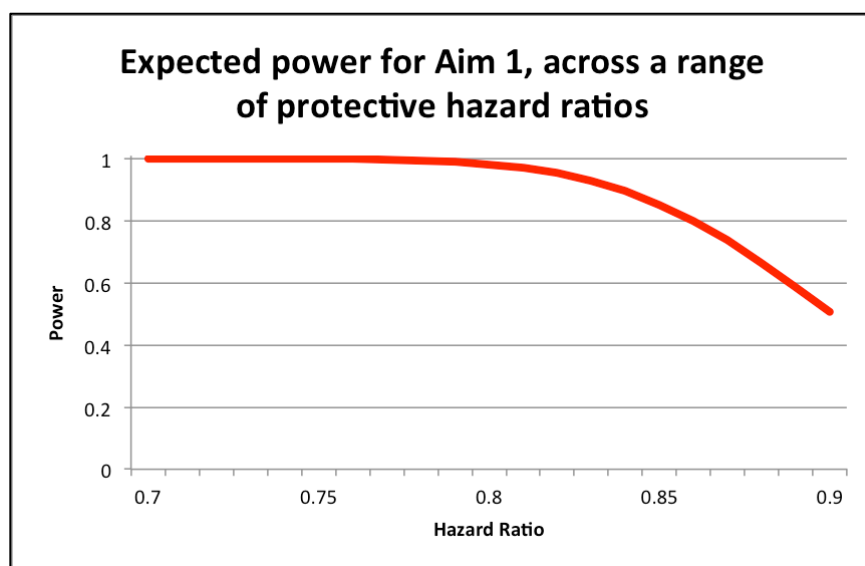


Figure 4. Expected statistical power to explore the association between school enrollment and teen pregnancy over a range of hazard ratios (0.7 to 0.9)

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

Table 3. The expected confidence limits and statistical power across a range of protective hazard ratios for the association between school enrollment and pregnancy

HR	0.50	0.60	0.70	0.80	0.90
95% CI	(0.45, 0.56)	(0.54, 0.67)	(0.63, 0.78)	(0.72, 0.89)	(0.81, 1.00)

C.4 Aim 2: Alcohol outlet visits and sexual risk behavior

C.4.1 Study design

To explore the association between alcohol outlet visits and sexual risk behavior, we will conduct a cross-sectional analysis using data collected in the baseline quantitative survey for HPTN 068. Information on the exposure (number of visits to a tavern in the last six months) and outcomes (sexual debut and unprotected sex in the last three months) will be collected at the individual-level at the same point in time. All young women with baseline data collected in the parent study will be included in the analysis.

C.4.2 Data collection and variable measurement

Outcome assessment: The outcome we are exploring in Aim 2 is sexual risk behavior. We will define sexual risk behavior using two variables self-reported in the HPTN 068 baseline questionnaire:

- 1) **Sexual debut:** This binary variable will be defined by the self-reported answer to the question: "Have you ever had vaginal sex?" (yes/no)
- 2) **Unprotected sex acts in the last three months:** This variable will be a count variable defined by subtracting the total number of condom-protected vaginal sex acts from the total number of vaginal sex acts reported in the last three months. We will consider this variable in two ways: using both the full sample and a sample restricting to those who reported at least one sex act in the last three months. For the full sample, those who reported no sex acts at all will receive a zero value for unprotected sex acts.

To minimize the bias that may come from answering personal questions about sexual behavior to an interviewer, an ACASI (audio computer-assisted interviewing) component was incorporated into the HPTN 068 baseline questionnaire. The ACASI component allows participants to privately read or listen to audiotaped questions and log their responses directly in the computer, without having direct interaction with the interviewer. To help prevent data entry errors, skip patterns and warnings for implausible or impossible responses were coded into the program.

The frequencies of sexual debut and descriptive statistics for unprotected sex are shown in Table 4. In the full sample, about 27% of young women reported experiencing sexual debut before baseline and 8% reported having at least one unprotected sex act in the last three months. When restricting to the population who reported any sex acts in the last three months (n=528), the proportion reporting unprotected sex rose to 36%. In the analysis for unprotected sex, we will carefully explore whether our results are sensitive to sample restriction. If results are not sensitive to sample restriction, we will use the full sample to maximize statistical efficiency.

Table 4. Frequency of sexual debut and descriptive statistics for unprotected sex acts in baseline sample

Outcome	N	%			
Sexual debut	673	26.6			
	Mean	Median	SD	Range	# of zero values (%)
Unprotected sex acts ¹	0.4	0	4.9	0-140	2344 (92.4)
Unprotected sex acts ²	2.1	0	10.7	0-140	336 (63.6)

¹Descriptive statistics from the full sample (n=2536)

²Descriptive statistics restricting to the sample of those reporting sex acts in the last three months (n=528)

Figure 5 shows the univariate distribution of unprotected sex acts, restricting to the sample of young women who reported any sex acts in the last three months ($n=528$). Although count data like these are often Poisson distributed, we see that the Poisson distribution is not a good fit to the data, likely due to the large numbers of zero responses. The univariate distribution of the full sample is even more positively skewed as all the young women excluded from the restricted sample because they have not had sex in the last three months, by definition, have zero unprotected sex acts.

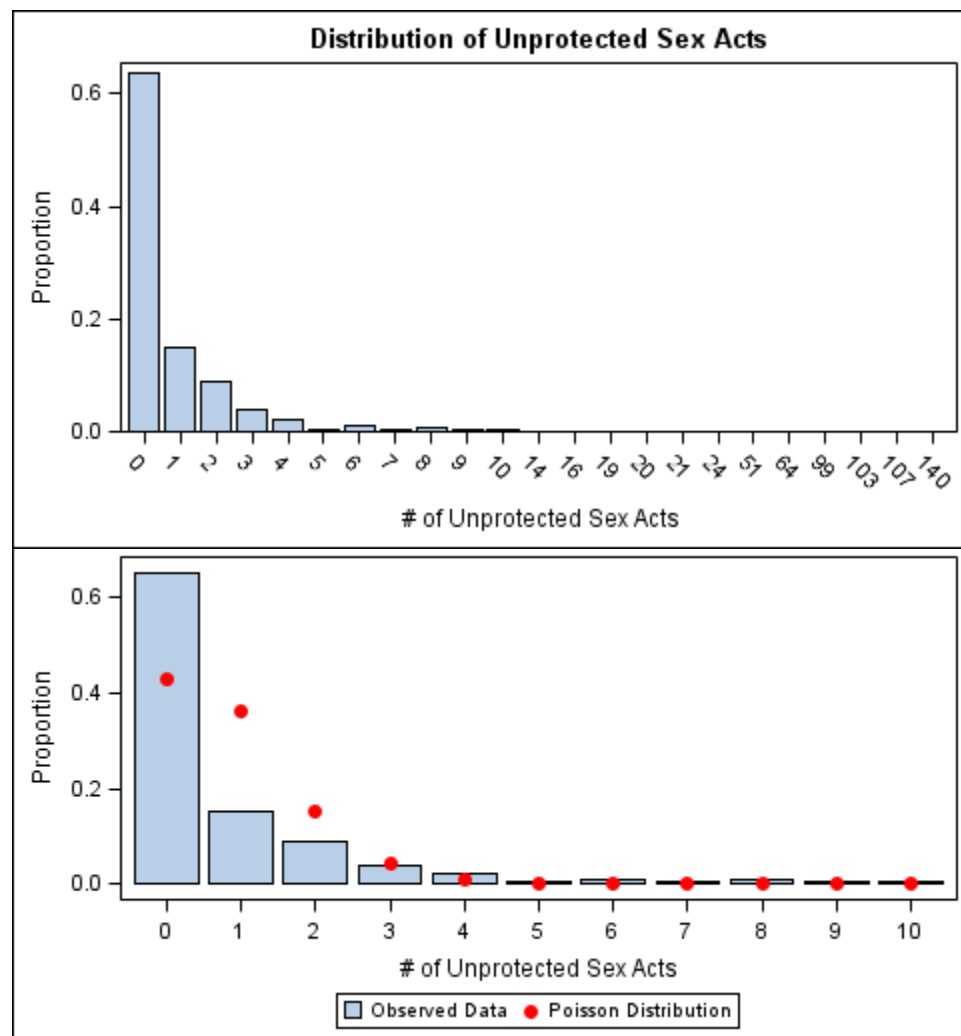


Figure 5. Histogram of unprotected sex acts among respondents reporting sex in last three months ($n=528$) across the full range of responses (top), and restricted to 0 to 10 visits with a Poisson distribution overlaid (bottom)

Exposure assessment: The exposure we are assessing in Aim 2 is frequency of tavern visits. This variable will be defined as the number of self-reported visits to a tavern in the last six months. This is a continuous count variable and we will assess the functional form of it's relationship with both outcome variables by considering different categorizations and transformations. Table 5 shows descriptive statistics for the number of tavern visits and Figure 6 shows a histogram of the univariate distribution. Overall, in the baseline sample, the mean number of tavern visits in the last six months was 1.6 with a standard deviation of 10.7. The values ranged from 0 to 479. The distribution is positively skewed with a large number of zero responses and the majority of responses ranging between zero and ten. In the

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

analysis, we will examine the sensitivity of our results to the removal of potential outliers with implausibly high responses.

Table 5. Descriptive statistics for number of tavern visits in the last six months

Exposure	Mean	Median	SD	Range
# Visits to tavern	1.6	0	10.7	0-479

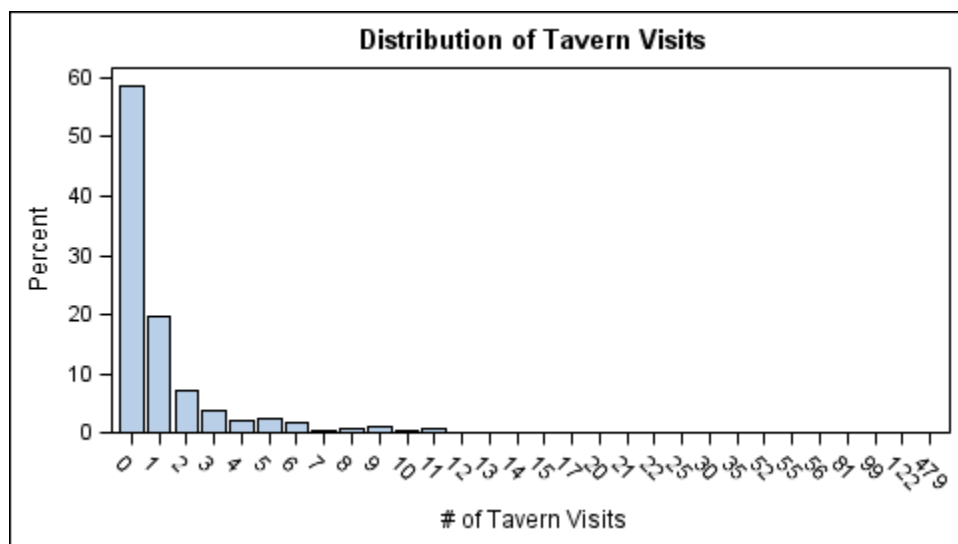


Figure 6. Histogram of number of tavern visits in the last six months in the full sample at baseline

Covariate assessment:

We will explore the influence of several key covariates on the relationship between activity participation and sexual risk behavior. Specifically we will examine:

- 1) **Age**, defined as age, in years, at the time of baseline interview. We hypothesize that increasing age will be associated with an increased number of tavern visits and increased likelihood of sexual debut.
- 2) **Education**, defined as the current grade in which the young woman is enrolled at baseline (an eligibility criterion for HPTN 068 is current school enrollment at baseline). We hypothesize that increasing education will be associated with a decreased number of tavern visits and decreased likelihood of sexual debut.
- 3) **Household size**, defined as the total number of individuals sharing a household with the young woman. We hypothesize that increasing household size will be associated with an increased number of tavern visits and increased likelihood of sexual debut.
- 4) **Primary caregiver relationship**, defined as how the young woman is related to her primary caregiver: daughter, sibling, niece, grandchild, other. We hypothesize that having a primary caregiver relationship other than daughter will be associated with an increased number of tavern visits and an increased likelihood of sexual debut.
- 5) **Village**, defined as the village of residence of the young woman at baseline. We hypothesize that villages will have different levels of alcohol outlet accessibility with different opportunities for tavern visits to occur. We hypothesize that village will be associated with sexual debut only indirectly through the association between village and household SES.
- 6) **Household SES**, defined by a continuous index measuring per capita monthly household consumption of food and durable goods. We hypothesize that an increased household SES will

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

be associated with number of tavern visits though the direction of the association may be positive (increased SES leads to more expendable income to be spent at taverns) or negative (increased SES associated with decreased visits to taverns through decreased household size and increased parental monitoring). We hypothesize that an increased household SES will be associated with a decreased likelihood of sexual debut.

The hypothesized relationships between these key covariates, the exposure (visits to taverns), and the primary outcome (sexual debut) are depicted in Figure 7 below.

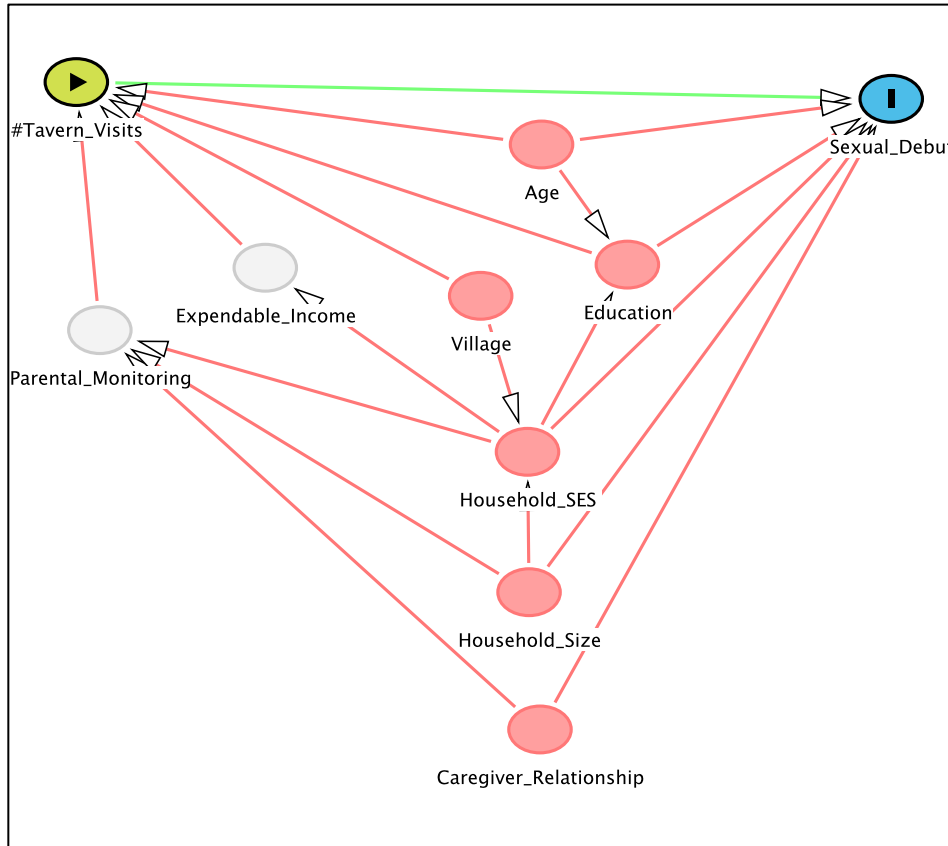


Figure 7. Directed acyclic graph depicting the hypothesized relationship between visits to taverns and sexual debut.

C.4.3 Statistical approach

We will use a logistic regression model to estimate the association between number of tavern visits and sexual debut (Equation 4.1). This approach is appropriate for the dichotomous sexual debut outcome we are investigating.

$$\text{Equation 4.1: } P(\text{Debut} = 1) = \frac{1}{1 + \exp^{-\alpha + \beta_1(\text{tavernvisit}) + \beta_2 X_2}}$$

where X is a matrix of the potential covariates in our final adjustment set.

We will use a zero-inflated Poisson model to estimate the association between number of tavern visits and number of unprotected sex acts (Equation 4.2).⁽¹³¹⁾ This approach is appropriate given the probability of reporting zero unprotected sex acts in excess of the Poisson distribution (see Figure 5). If the unprotected sex act data are overdispersed (the variance is greater than the mean)

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

even after using a zero-inflated model, a Poisson distribution may not be appropriate. We will test for overdispersion by examining the model output of the zero-inflated Poisson model. If the deviance divided by the degrees of freedom is substantially greater than one, then we will consider a zero-inflated negative binomial model. A negative binomial model estimates a dispersion parameter to create a more flexible model that allows the mean to differ from the variance. A likelihood ratio test will be performed assess whether the inclusion of this additional parameter significantly improves model fit over the Poisson model.

$$\text{Equation 4.2: } P(\#SexActs = j) = \begin{cases} \pi_i + (1 - \pi_i)\exp^{-\lambda_i} & \text{if } j = 0 \\ (1 - \pi_i)\frac{\lambda_i^j \exp^{-\lambda_i}}{j!} & \text{if } j > 0 \end{cases}$$

where π_i is the probability of zero unprotected sex acts in excess of the underlying Poisson distribution for individual i ; λ_i is the expected unprotected sex act count for individual i as estimated by the model specified as $\alpha + \beta_1(tavernvisit) + \beta_2X_2$.

We will assess for the potential confounding effects of age, education, village, SES, household size, and caregiver relationship. First, we will identify a minimally sufficient adjustment set from the directed acyclic graph (Figure 7). Only covariates in the minimally sufficient adjustment set that alter the effect estimate by greater than 10% will be retained in the final adjusted models. No *a priori* effect measure modifiers will be explored.

C.4.4 Power

We parameterized our power calculation for a simple logistic regression model estimating the odds ratio for a one-unit increase in number of tavern visits on sexual debut. We assumed a two-sided test for the statistical significance of the likelihood ratio chi-square statistic at $\alpha=0.05$ and a total sample size of **2536**. We further assumed that the tavern visit count variable was Poisson distributed with a mean of 1.6, as calculated from the baseline survey data. We assumed that the proportion of young women who had sexually debuted among those reporting no visits to taverns in the last six months was 24%, as calculated from the baseline survey data. Finally, we varied the expected odds ratios from 1.05 to 1.25 to cover a range of plausibly small effect estimates in the hypothesized direction.

Figure 8 shows the power curve over the range of odds ratios above the null. Overall, for all odds ratios greater than 1.11 we maintain a statistical power above 80% to estimate the effect of a one unit increase in the number of visits to taverns on odds of sexual debut.

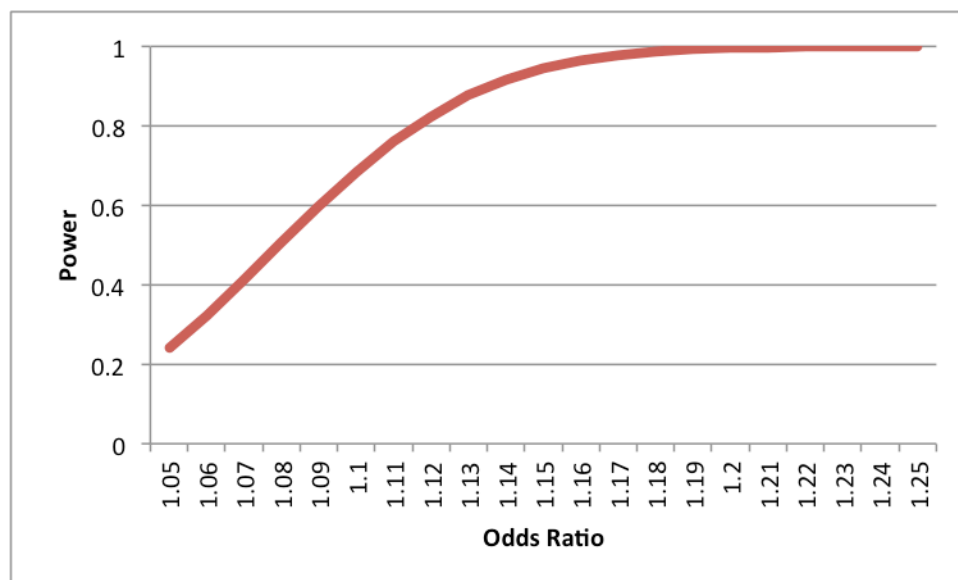


Figure 8. Expected statistical power to explore the association between number of tavern visits and sexual debut over a range of odds ratios from 1.05 to 1.25

C.5 Aim 3: Alcohol outlet density and sexual risk

C.5.1 Study design

To explore the influence of alcohol outlet density on sexual risk as measured by prevalent HSV-2 infection, we will perform a cross-sectional study using biologic data collected at baseline in HPTN 068, geographic data collected and maintained by the Agincourt HDSS, and village asset data collected as part of a community survey performed as a complementary study to HPTN 068. Alcohol outlet density will be measured and analyzed at the village-level while the sexual risk outcome will be measured and analyzed at the individual-level. The design and analysis of this aim is grounded in a multilevel framework because we are including variables at both the individual- and village-level.

C.5.2 Data collection and variable measurement

Outcome assessment:

The primary outcome of interest in Aim 3 is **prevalent herpes simplex virus 2 (HSV-2) infection**. There were a total of 118 HSV-2 infections (4.7%) identified in the study population at baseline. The herpes virus is a sexually transmitted virus, like HIV, and can therefore be considered as an indicator of sexual risk. Every HPTN 068 participant is tested for HSV-2 infection at baseline, prior to randomization. HSV-2 testing is performed with a Kalon assay (cut-off of 1.5). We will use the same definitions for prevalent HSV-2 infection in this ancillary study.

Although data on prevalent HIV infection were also collected by HPTN 068, we are not investigating it as a sexual risk outcome for two major reasons. First, HIV is typically less prevalent than HSV-2, as confirmed by the 81 prevalent HIV infections (3.2%) identified in the study sample at baseline. It is worth noting that this study population is young (mean age: 15.6), which may account for the relatively low prevalence. It is expected that HIV infections in the study population will rise as the cohort ages; however, for the purposes of this proposed study, the HIV prevalence may be too low to efficiently investigate our research questions. Second, some of these HIV infections may not be the result of sexual risk behavior but instead may be the result of long-term survival after vertical transmission or infection due to childhood sexual abuse. As this study aims to investigate the social and structural determinants

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

of sexual behavior, vertical infections or childhood sexual abuse-related infections would not be informative. A benefit of using HSV-2 as the primary outcome is that, unlike HIV, HSV-2 infections in young women cannot be due to perinatal transmission.

Exposure assessment:

The exposure of interest for this aim is **alcohol outlet density**. For each of 24 villages from which young women in HPTN 068 were enrolled, we will calculate the alcohol outlet density by dividing the number of alcohol vendors in each village per square kilometer. Each young woman will be assigned the alcohol outlet density value of her village of residence.

The number of alcohol vendors per village was reported by consensus among several key informants from each village through a community mapping exercise undertaken in a HPTN 068 sub-study. The number of taverns and the number of bottle shops (liquor stores) were recorded separately in the mapping exercise. We will consider defining the exposure as the density of taverns and as the density of taverns and bottle shops combined. To maximize statistical efficiency, we will use the combined density exposure if tests of homogeneity indicate that effect estimates for the different exposure definitions are not statistically different from each other. The number of square kilometers in each village will be derived from the village outline shapefile in the geographic database created and maintained by the Agincourt HDSS using ArcGIS software. The preliminary density of alcohol outlets, including both taverns and bottle stores, ranges from 0 to 4.2 outlets per square kilometer. The geographic distribution of this variable is shown in Figure 9.

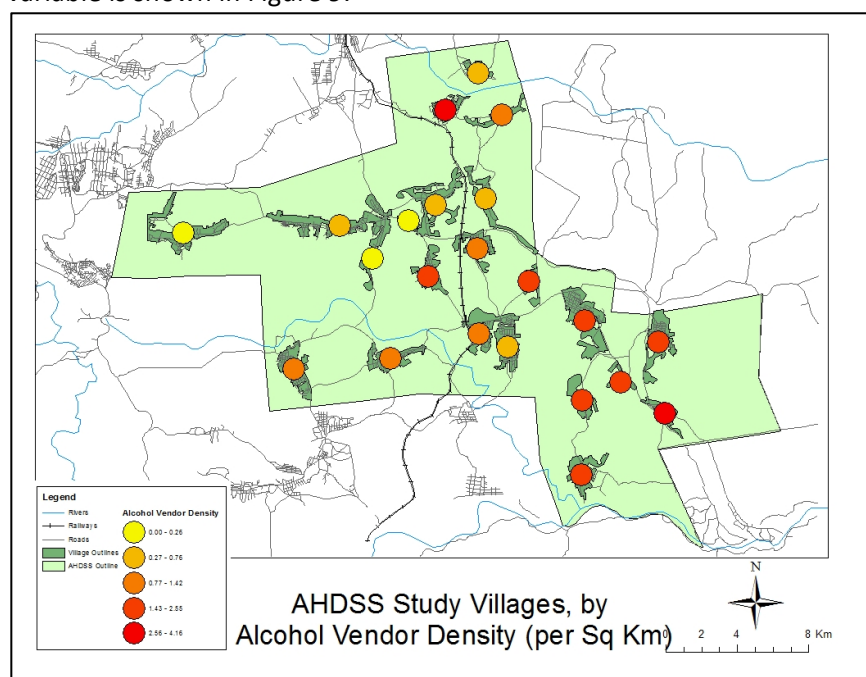


Figure 9. Geographic distribution of alcohol vendor density across villages in the Agincourt HDSS study site

The descriptive statistics at the individual level for number of taverns, number of bottle shops, and alcohol outlet density are listed in Table 6 and the histogram for the individual distribution of alcohol outlet density is shown in Figure 10. The mean alcohol outlet density in the full sample is 1.6 with a standard deviation of 0.9. The data appear to be fairly uniformly distributed across the range of values, though there may be slightly more data at moderate densities than for the lowest and highest densities.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

Table 6. Descriptive statistics for village alcohol outlets and alcohol outlet density at individual-level

Variable	Mean	Median	SD	Range	Skewness	Kurtosis
# Taverns	3.6	4	1.6	0-6	-0.51	-0.58
# Bottle Shops	0.6	0	0.8	0-2	0.77	-0.99
Alcohol outlet density ¹	1.6	1.4	0.9	0-4.2	0.44	0.03

¹Alcohol outlet density calculated considering both taverns and bottle shops as alcohol outlets

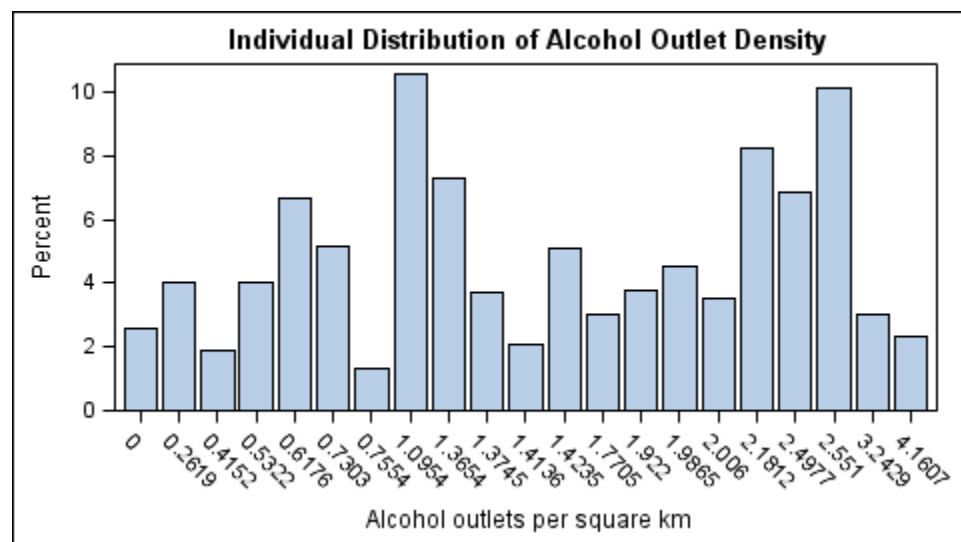


Figure 10. Histogram of the individual-level distribution of alcohol outlet density within home village

Covariate assessment:

We will also collect information on several key covariates to consider in the association between alcohol vendor density and sexual risk. Specifically, we will consider:

- 1) **Age**, defined as age in years at baseline. We hypothesize that increasing age will be associated with increased likelihood of HSV-2 infection but not directly associated with alcohol outlet density.
- 2) **Number of visits to taverns**, defined as the number of self-reported visits to a tavern in the last six months, as recorded in the baseline survey. We hypothesize that this variable acts as a mediator between alcohol outlet density and HSV-2 infection.
- 3) **Household SES**, defined by a continuous index measuring per capita monthly household consumption of food and durable goods collected in the household baseline questionnaire asked of the primary caregiver of each study participant. We hypothesize that increasing household SES is associated with lower likelihood of HSV-2 infection and associated with alcohol outlet density through mean village-level SES.
- 4) **Education**, defined as the current grade of school in which the young woman is enrolled at baseline. We hypothesize that increasing education is associated with a lower likelihood of HSV-2 infection. We do not hypothesize that education is directly associated with alcohol outlet density.
- 5) **Mean village-level SES**, defined as the mean SES for all households in the village, measured as an index of household assets in the HDSS census in 2007. We hypothesize that increasing mean village-level SES is associated with an increase in alcohol outlet density. We hypothesize that increasing mean village-level SES is associated with HSV-2 infection through its association with household level SES.

Dissertation Proposal: Molly Rosenberg
School Enrollment, Alcohol Outlets, and Sexual Risk

- 6) **Village population**, defined as the most recent population estimate for each village, as derived from the HDSS census. We hypothesize that increasing village population is associated with increased alcohol outlet density and increased likelihood of HSV-2 infection.

The hypothesized relationships between these covariates, the alcohol outlet density exposure and the primary outcome (HSV-2 infection) are depicted in the directed acyclic graph in Figure 11.

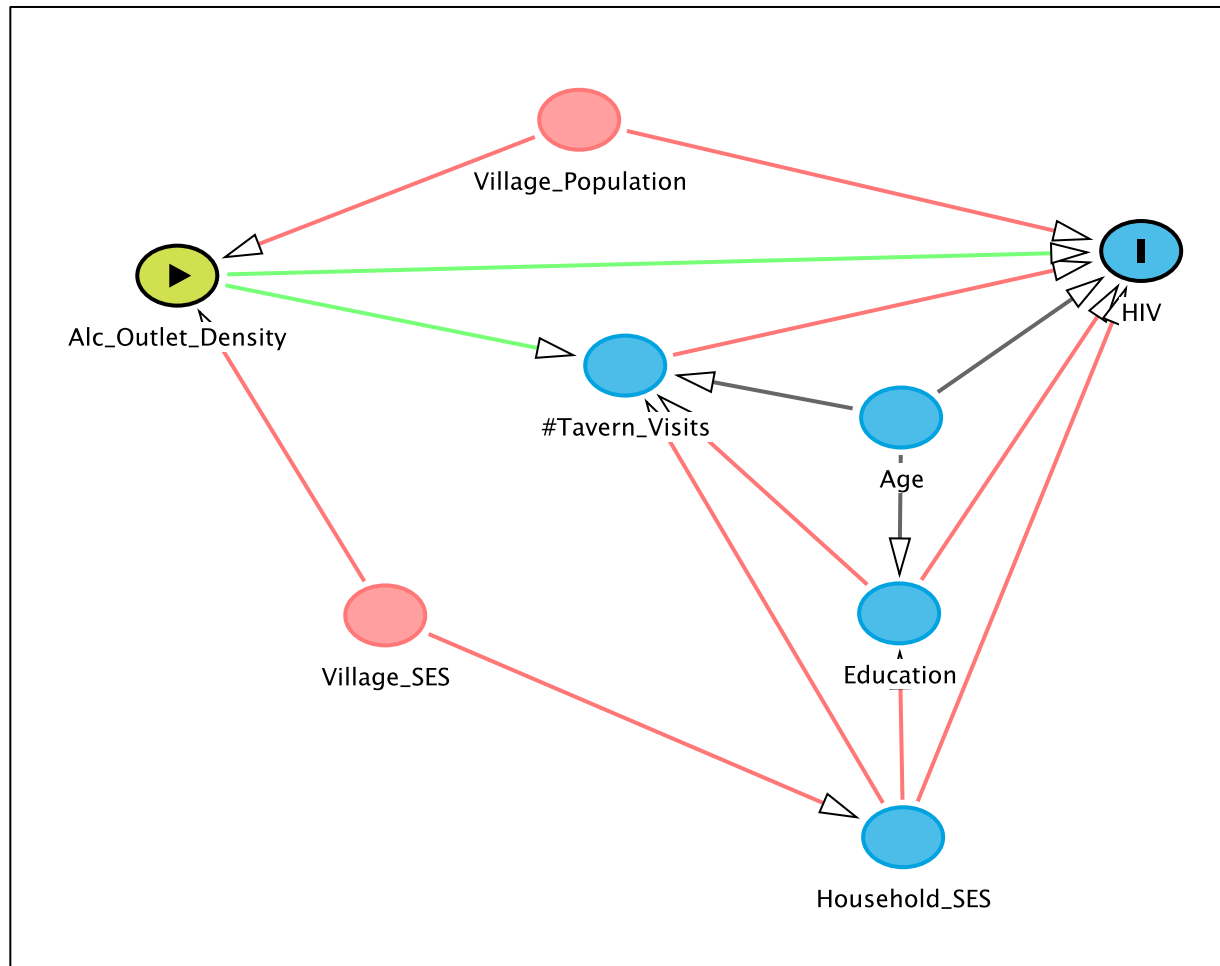


Figure 11. Directed acyclic graph depicting the hypothesized relationship between alcohol outlet density and HSV-2 infection.

C.5.3 Statistical approach

As noted previously, the statistical analysis for Aim 3 will be grounded in a *Neighborhoods and Health* theoretical framework. This framework was developed to explore the health effects of factors at multiple levels of organization.(32) Here, we are simultaneously examining the effects of factors at the individual- and the village-level so a *multilevel analysis* will be performed to account for the non-independent nature of the data.(132, 133)

The hierarchical model will be built in two stages. First, an individual-level logistic regression model will be constructed, indexed by individual and village, predicting the likelihood of individual HSV-2 infection based on the individual-level factors for which we wish to control (Equation 5.1).

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

Level 1 model

Equation 5.1: $HSV_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + e_{ij}$

(i and j reference each individual and village, respectively; X represents a matrix of all individual covariates in the adjustment set; and e represents the error)

Next, the parameters defined in the individual (or ‘Level 1’) regression model will be modeled as functions of the village-level in ‘Level 2’ equations (Equations 5.2-5.3). Here, we include our village-level exposure of interest (alcohol outlet density) as well as any village-level covariates in the model predicting the intercept term of our Level 1 model. In this simple example, we do not allow the effects of the individual-level covariates to vary by village, though this complexity could be incorporated by including a random error term in Equation 5.3. Even further complexity could be incorporated by allowing for interactions between Level 1 and Level 2 variables. We will weigh the potential benefits of these more complex model specifications with likelihood ratio tests comparing models with and without random error terms and with and without higher-level interactions.

Level 2 models

Equation 5.2: $\beta_{0j} = \gamma_{00} + \gamma_{10}AOD + u_{0j}$

Equation 5.3: $\beta_{1j} = \gamma_{01}$

(AOD represents the value of the alcohol outlet density variable; u represents the random error at the group level)

The final multilevel logistic model will plug equations 5.2-5.3 back into the model presented in Equation 5.1 to include a fixed portion that remains true across all villages and a random portion that allows the effect to vary between villages (Equation 5.4). Exponentiating the gamma coefficient for the alcohol outlet density parameter will provide an odds ratio of the average effect of the exposure on odds of HSV-2 infection across all villages, adjusting for individual-level confounders (and other village-level covariates, if they are included). The standard errors estimated for this effect estimate are adjusted for non-independence because of the multilevel specification of the model.

Final multilevel model

Equation 5.4: $HSV_{ij} = \gamma_{00} + \gamma_{10}AOD + \gamma_{01}X_{1ij} + u_{0j}$

We will assess for the potential confounding effects of age, educational attainment, household SES, village SES, and village population. First we will identify a minimally sufficient adjustment set from a directed acyclic graph (Figure 11). Only covariates in the adjustment set that alter the effect estimate by greater than 10% will be retained in the final model. No *a priori* potential effect measure modifiers will be explored.

If the association between alcohol outlet density and HSV-2 is statistically significant, we wish to assess whether whether or not the association is mediated by individual visits to taverns. To do this we will add the individual tavern visit variable to the model, adjusting the minimally sufficient adjustment set updated for the bias induced by collider stratification. If the effect remains significant (at $\alpha=0.05$) after controlling for individual tavern visits, we will conclude that the contextual effect is significant. All statistical analyses for this aim will be performed using R statistical software.

Multilevel modeling is an appropriate tool to use to answer these research questions for two major reasons. First, it correctly adjusts the standard error of effect estimates to account for the fact that young women living in the same village likely have correlated outcomes because they are exposed to the same environment. Second, the multilevel framework allows us to create a model at the level of

Dissertation Proposal: Molly Rosenberg School Enrollment, Alcohol Outlets, and Sexual Risk

complexity the data suggest and can support. For example, we will explore how, if at all, the effect of individual covariates on sexual risk varies across villages.

C.5.6 Power

To account for the inflated precision that comes with the use of clustered data, we need to adjust our projected sample size by a factor equal to the *design effect* to accurately estimate power. We calculate the *design effect* in this study to be **1.84** using the following formula:

$$1+(n-1)\rho \quad (134)$$

where n = average participants per village = 85; and ρ = the intracluster correlation coefficient, (ICC), which we estimated to be 0.01 based on the fact that ICC typically falls between 0.01 and 0.02 in human studies.(135)

We parameterized our power calculation for a simple logistic regression model estimating the odds ratio for a one-unit increase in alcohol outlet density on the odds of HSV-2 infection. Given the design effect of 1.84, our effective sample size is decreased from 2536 to **1378**. We further parameterized our power calculation assuming a total proportion of **HSV-2 outcomes in the unexposed at 4.7%**. We assumed that the alcohol outlet density variable was normally distributed with a mean of 1.6 and standard deviation of 0.9 as calculated from the baseline survey data. Finally, we varied the expected magnitude of effect from an odds ratio of 1.1 to 2.0. The result of the power calculation assuming a two-sided test for the statistical significance of the likelihood ratio chi-square statistic at $\alpha=0.05$ is shown in Figure 12. Although, as expected, this aim is underpowered at odds ratios close to the null, we maintain a statistical power above 80% for all odds ratios above 1.50.

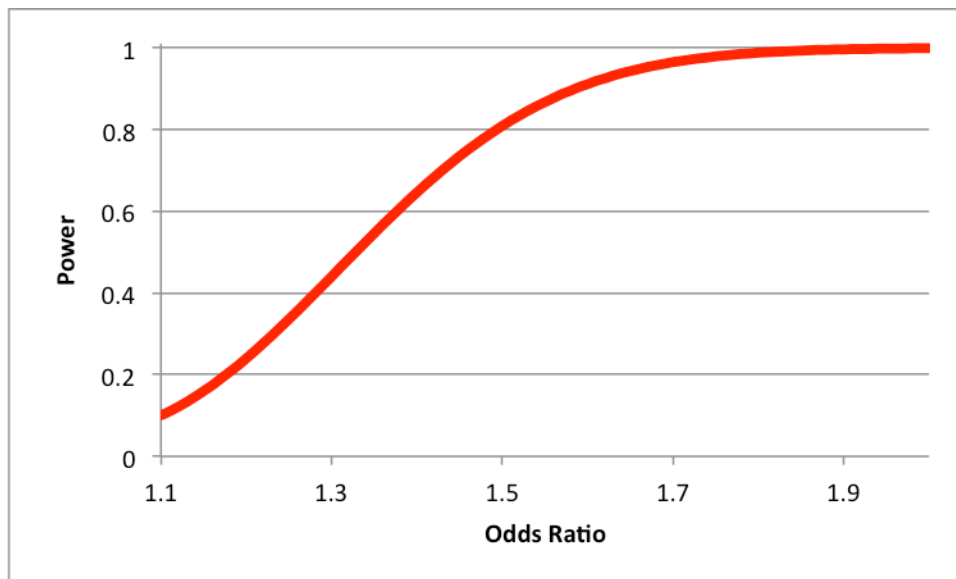


Figure 12. Expected statistical power to explore the association between alcohol outlet density and HSV-2 across a range of odds ratios from 1.1 to 2.0

D. STRENGTHS AND LIMITATIONS

This proposed study offers an important contribution to the epidemiologic literature on adolescent sexual risk for several reasons. All three aims are designed to explore the potential influence of routine activities on sexual risk, an understudied set of structural exposures. Although the hypothesis that

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

routine activities may influence sexual risk is well-supported theoretically, these potential exposures have been largely ignored in the epidemiologic literature to date. Another important aspect of these exposures is their inherent link to practical interventions. Unlike some demographic or individual-level factors that are non-modifiable, if protective or risky associations are discovered with certain routine activities, this research could be easily extended into the development of analogous preventive interventions (e.g. decrease the costs of school attendance; target safe sex messages and condom distribution in taverns).

The datasets employed in this study are appropriate to explore the associations between routine activities and sexual risk. All three aims of the proposed study make use of existing, rich datasets that have well-measured indicators of routine exposures and sexual risk outcomes. This will allow us to efficiently answer our research questions with relatively little investment of time or money. In Aim 1, we use information from the Agincourt HDSS census to explore the association between school enrolment and teen pregnancy. In Aims 2-3, we use information from baseline data collection activities for HPTN 068 to explore: 1) the association between visits to alcohol outlets and sexual risk behavior, and 2) the association between village alcohol outlet density and HSV-2 prevalence.

We will perform our analyses on data from a study population of young women living in rural South Africa. Our research questions are particularly pertinent to this study population because, as mentioned previously, they are at extraordinarily high risk for negative sexual risk outcomes such as HIV(3) and teenage pregnancy.(9) It is important to note, however, that the results of this study may not be generalizable to either young men or to adolescents in other parts of the world. It is also important to note that some of the eligibility criteria for HPTN 068 (enrolled in school, live with parent/guardian, have a bank account) may further limit the generalizability of the results outside of our study population for Aims 2 and 3. However, as most young women are enrolled in school in South Africa, particularly those in younger ages, few were excluded from the study for these reasons. We plan to conduct an informal sensitivity analysis around these selection criteria by exploring the differences between HPTN 068 study participants and all young women in the same age range on key demographic data from the HDSS census (e.g. age, SES, education, family structure).

The study design employed for each aim is appropriate given the data structure and the specific research question. The longitudinal design of Aim 1 is advantageous for several reasons. First, it will allow us to meaningfully incorporate time into the analysis. This added information will allow us to know *when* the pregnancy occurred as opposed to simply knowing whether it happened or not. This is important because, with an outcome like teen pregnancy, the timing of the pregnancy in adolescence may have implications for how disruptive it may be to the life course of the adolescent and how prepared she may be for motherhood. Second, the longitudinal design will allow us to establish temporality between the exposure and the outcome. This is an important step to be able to make inferences about the influence of school enrollment on pregnancy, particularly given that the reverse relationship is likely true as well (pregnancy at time one likely influences future school enrollment status at time two).(56)

The analyses for Aims 2 and 3 are both cross-sectional. Two general weaknesses of cross-sectional studies are the uncertainty regarding the direction of association between exposure and outcome and the potential bias induced using prevalent instead of incident outcomes. These concerns are mitigated for Aim 2 for the following reasons:

- 1) This analysis aims to establish whether or not there is an association between visits to alcohol outlets and sexual behavior. If a significant association is found, the investment in future

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

longitudinal studies to establish the temporal relationship between activities and sexual behavior may be worthwhile.

- 2) Although we hypothesize that visits to taverns may influence sexual behavior, if the opposite is true (sexual behavior influences the frequency of tavern visits), the information is still valuable. For example, the knowledge that groups of high-risk young women gather in certain places (like taverns) could be useful for planning preventive interventions as efficient place-based recruitment of potential participants or place-based delivery of information/materials could be performed.
- 3) At baseline, this is a cohort of relatively young women (mean age: 15.6), most of whom have not yet experienced sexual debut (proportion with sexual debut: 26.6%) or only very recently sexually debuted (median age at debut: 16). Therefore, we assume that the duration of the prevalent sexual behavior outcomes used in this analysis is fairly short and that they reasonably approximate incident outcomes, minimizing the potential for prevalence-incidence bias. Also, the nature of the outcome we are exploring (sexual behavior) does not naturally lend itself to some of the major problems with using prevalent versus incident outcomes. As mortality is not immediately associated with the sexual behavior outcomes we are exploring, it is unlikely that there will be young women who did not survive the short duration between the time that they first experience a risky sexual behavior and the baseline data collection.

The above concerns about cross-sectional designs are mitigated for Aim 3 for the following reasons:

- 1) As in Aim 2, establishing a causal relationship between the exposure and outcome is not an objective of this study. The aim of this analysis is to assess whether alcohol outlet density is associated with sexual risk.
- 2) Alcohol outlet density exposure can be assumed to be relatively constant over recent history as the opening or closing of a business is a rather long process and the influence of changes in a single outlet's status is not likely to influence the overall density measure in a village strongly. However, the sexual transmission of HSV-2 is assumed to be relatively recent in this cohort of young women because of their young age. Therefore directionality between the exposure and the outcome can be better inferred.
- 3) Because the difference between the date of sexual transmission of HSV-2 and the date of baseline data collection is likely small due to the young age of this cohort, the potential for prevalence-incidence bias is minimized. It is unlikely there are many young women who died from fast progression of disease before they had a chance to enroll in the study because HSV-2 does not progress quickly.(137)

Finally, Aim 3 of this study employs a novel set of methodological tools to explore the association between alcohol outlet density and sexual risk. Multilevel models are still relatively rare in epidemiologic research, though there is considerable interest in exploring questions that would best be served using multilevel methods. The successful use of this approach in the proposed study would be a step towards moving the field toward broader adoption of these promising methods.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

E. LITERATURE CITED

- 1 UNAIDS. UNAIDS report on the global AIDS epidemic; 2012.
- 2 UNAIDS. South Africa country overview. Geneva: UNAIDS; 2011.
- 3 Pettifor AE, Rees HV, Kleinschmidt I, et al. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *AIDS*. 2005 Sep 23;**19**(14):1525-34.
- 4 UNAIDS. Global AIDS Response Progress Report, 2012: Republic of South Africa; 2012.
- 5 Ruedinger E, Cox JE. Adolescent childbearing: consequences and interventions. *Curr Opin Pediatr*. 2012 Aug;**24**(4):446-52.
- 6 Black AY, Fleming NA, Rome ES. Pregnancy in adolescents. *Adolesc Med State Art Rev*. 2012 Apr;**23**(1):123-38, xi.
- 7 Pinzon JL, Jones VF. Care of adolescent parents and their children. *Pediatrics*. 2012 Dec;**130**(6):e1743-56.
- 8 WHO. Fact Sheet: Adolescent Pregnancy; 2012.
- 9 United Nations Statistics Division. Live births by age of mother and sex of child, general and age-specific fertility rates: latest available year, 2000–2009; 2000-2009.
- 10 Republic of South Africa. South Africa Demographic and Health Survey; 2003.
- 11 Serwadda D, Wawer MJ, Musgrave SD, Sewankambo NK, Kaplan JE, Gray RH. HIV risk factors in three geographic strata of rural Rakai District, Uganda. *AIDS*. 1992 Sep;**6**(9):983-9.
- 12 UNAIDS. Women out loud: How women living with HIV will help the world end AIDS; 2012.
- 13 Wojcicki JM. Socioeconomic status as a risk factor for HIV infection in women in East, Central and Southern Africa: a systematic review. *J Biosoc Sci*. 2005 Jan;**37**(1):1-36.
- 14 Pettifor AE, van der Straten A, Dunbar MS, Shiboski SC, Padian NS. Early age of first sex: a risk factor for HIV infection among women in Zimbabwe. *AIDS*. 2004 Jul 2;**18**(10):1435-42.
- 15 Parazzini F, Cavalieri D'oro L, Naldi L, et al. Number of sexual partners, condom use and risk of human immunodeficiency virus infection. *Int J Epidemiol*. 1995 Dec;**24**(6):1197-203.
- 16 Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. *Sex Transm Infect*. 1999 Feb;**75**(1):3-17.
- 17 Vundule C, Maforah F, Jewkes R, Jordaan E. Risk factors for teenage pregnancy among sexually active black adolescents in Cape Town. A case control study. *S Afr Med J*. 2001 Jan;**91**(1):73-80.
- 18 Lavin C, Cox JE. Teen pregnancy prevention: current perspectives. *Curr Opin Pediatr*. 2012 Aug;**24**(4):462-9.
- 19 Pettifor AE, Levandowski BA, MacPhail C, Padian NS, Cohen MS, Rees HV. Keep them in school: the importance of education as a protective factor against HIV infection among young South African women. *Int J Epidemiol*. 2008 Dec;**37**(6):1266-73.
- 20 Lopman B, Lewis J, Nyamukapa C, Mushati P, Chandiwana S, Gregson S. HIV incidence and poverty in Manicaland, Zimbabwe: is HIV becoming a disease of the poor? *AIDS*. 2007 Nov;**21** Suppl 7:S57-66.
- 21 McHunu G, Peltzer K, Tutshana B, Seutlwadi L. Adolescent pregnancy and associated factors in South African youth. *Afr Health Sci*. 2012 Dec;**12**(4):426-34.
- 22 Blum R, Mmari K. Risk and protective factors affecting adolescent reproductive health in developing countries: WHO; 2004.
- 23 Hirschi T, Stark R. Hellfire and delinquency. *Social Problems*. 1969:202-13.
- 24 Hirschi T. Causes of delinquency: Transaction publishers; 2002.
- 25 Cohen LE, Felson M. Social change and crime rate trends: A routine activity approach. *American sociological review*. 1979:588-608.
- 26 Felson M, Gottfredson M. Social indicators of adolescent activities near peers and parents. *Journal of Marriage and the Family*. 1984:709-14.
- 27 Hawdon JE. Daily Routines and Crime Using Routine Activities as Measures of Hirschi's Involvement. *Youth & Society*. 1999;**30**(4):395-415.
- 28 Osgood DW, Wilson JK, O'Malley PM, Bachman JG, Johnston LD. Routine activities and individual deviant behavior. *American Sociological Review*. 1996:635-55.
- 29 Barnes GM, Hoffman JH, Welte JW, Farrell MP, Dintcheff BA. Adolescents' time use: Effects on substance use, delinquency and sexual activity. *Journal of Youth and Adolescence*. 2007;**36**(5):697-710.
- 30 Bandura A. Social foundations of thought and action: A social cognitive theory. 1986.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

- 31 Bandura A. Social cognitive theory and exercise of control over HIV infection. *Preventing AIDS: Theories and methods of behavioral interventions*. 1994:25-59.
- 32 Kawachi I, Berkman L, (Eds.). *Neighborhoods and Health*. New York, NY: Oxford University Press; 2003.
- 33 Kawachi I, Subramanian SV, Almeida-Filho N. A glossary for health inequalities. *J Epidemiol Community Health*. 2002 Sep;**56**(9):647-52.
- 34 UNESCO Institute for Statistics. Enrolment ratios by ISCED level, South Africa; 2009.
- 35 Strassburg S, Meny-Gibert S, Russell B. *Left unfinished: Temporary absence and drop-out from South African schools*. Johannesburg, South Africa: Social Surveys Africa; 2010.
- 36 South African Ministry of Basic Education. *School calendars for public schools*; 2010.
- 37 Wittenberg M. *The school day in South Africa*. Southern Africa Labour and Development Research Unit, University of Cape Town: SALDRU/CSSR Working Papers; 2005.
- 38 Modisaotile BM. The failing standard of basic education in South Africa. *Policy Brief*. 2012;**72**.
- 39 Mbuyisi M. *South Africa: A fateful decision*; 2009.
- 40 Doyle AM, Mavedzenge SN, Plummer ML, Ross DA. The sexual behaviour of adolescents in sub-Saharan Africa: patterns and trends from national surveys. *Trop Med Int Health*. 2012 Jul;**17**(7):796-807.
- 41 Hargreaves JR, Slaymaker E, Fearon E, Howe LD. Changes over time in sexual behaviour among young people with different levels of educational attainment in Tanzania. *J Int AIDS Soc*. 2012;**15 Suppl 1**:1-7.
- 42 Zuilkowski SS, Jukes MC. The impact of education on sexual behavior in sub-Saharan Africa: a review of the evidence. *AIDS Care*. 2012;**24**(5):562-76.
- 43 Barnighausen T, Hosegood V, Timaeus IM, Newell ML. The socioeconomic determinants of HIV incidence: evidence from a longitudinal, population-based study in rural South Africa. *AIDS*. 2007 Nov;**21 Suppl 7**:S29-38.
- 44 Bradley H, Bedada A, Brahmabhatt H, Kidanu A, Gillespie D, Tsui A. Educational attainment and HIV status among Ethiopian voluntary counseling and testing clients. *AIDS Behav*. 2007 Sep;**11**(5):736-42.
- 45 Hargreaves JR, Glynn JR. Educational attainment and HIV-1 infection in developing countries: a systematic review. *Trop Med Int Health*. 2002 Jun;**7**(6):489-98.
- 46 Hargreaves JR, Bonell CP, Boler T, et al. Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa. *AIDS*. 2008 Jan 30;**22**(3):403-14.
- 47 de Walque D, Nakiyingi-Miiró JS, Busingye J, Whitworth JA. Changing association between schooling levels and HIV-1 infection over 11 years in a rural population cohort in south-west Uganda. *Trop Med Int Health*. 2005 Oct;**10**(10):993-1001.
- 48 Michelo C, Sandoy IF, Fylkesnes K. Marked HIV prevalence declines in higher educated young people: evidence from population-based surveys (1995-2003) in Zambia. *AIDS*. 2006 Apr 24;**20**(7):1031-8.
- 49 Johnson LF, Dorrrington RE, Bradshaw D, du Plessis H, Makubalo L. The effect of educational attainment and other factors on HIV risk in South African women: results from antenatal surveillance, 2000-2005. *AIDS*. 2009 Jul 31;**23**(12):1583-8.
- 50 Jukes M, Simmons S, Bundy D. Education and vulnerability: the role of schools in protecting young women and girls from HIV in southern Africa. *AIDS*. 2008 Dec;**22 Suppl 4**:S41-56.
- 51 Hargreaves JR, Morison LA, Kim JC, et al. The association between school attendance, HIV infection and sexual behaviour among young people in rural South Africa. *J Epidemiol Community Health*. 2008 Feb;**62**(2):113-9.
- 52 Stroeken K, Remes P, De Koker P, Michielsen K, Van Vossolle A, Temmerman M. HIV among out-of-school youth in Eastern and Southern Africa: a review. *AIDS Care*. 2012;**24**(2):186-94.
- 53 Gavin L, Galavotti C, Dube H, et al. Factors associated with HIV infection in adolescent females in Zimbabwe. *J Adolesc Health*. 2006 Oct;**39**(4):596 e11-8.
- 54 Grant MJ, Hallman KK. Pregnancy-related school dropout and prior school performance in KwaZulu-Natal, South Africa. *Stud Fam Plann*. 2008 Dec;**39**(4):369-82.
- 55 Clark S, Mathur R. Dating, sex, and schooling in urban Kenya. *Stud Fam Plann*. 2012 Sep;**43**(3):161-74.
- 56 Marteleto L, Lam D, Ranchhod V. Sexual behavior, pregnancy, and schooling among young people in urban South Africa. *Studies in Family Planning*. 2008;**39**(4):351-68.
- 57 Duflo E, Dupas P, Kremer M, Sinei S. Education and HIV/AIDS prevention: evidence from a randomized evaluation in Western Kenya. *World Bank Policy Research Working Paper*. 2006(4024).
- 58 Baird S, Chirwa E, McIntosh C, Ozler B. The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women. *Health Econ*. 2010 Sep;**19 Suppl**:55-68.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

- 59 Manlove J. The influence of high school dropout and school disengagement on the risk of school-age pregnancy. *J Res Adolesc.* 1998;**8**(2):187-220.
- 60 Warner L, Newman DR, Austin HD, et al. Condom effectiveness for reducing transmission of gonorrhea and chlamydia: the importance of assessing partner infection status. *Am J Epidemiol.* 2004 Feb 1;**159**(3):242-51.
- 61 Niccolai LM, Rowhani-Rahbar A, Jenkins H, Green S, Dunne DW. Condom effectiveness for prevention of Chlamydia trachomatis infection. *Sex Transm Infect.* 2005 Aug;**81**(4):323-5.
- 62 Shlay JC, McClung MW, Patnaik JL, Douglas JM, Jr. Comparison of sexually transmitted disease prevalence by reported level of condom use among patients attending an urban sexually transmitted disease clinic. *Sex Transm Dis.* 2004 Mar;**31**(3):154-60.
- 63 Celentano DD. Editorial response: it's all in the measurement: consistent condom use is effective in preventing sexually transmitted infections. *Sex Transm Dis.* 2004 Mar;**31**(3):161-2.
- 64 Minnis AM, van der Straten A, Gerdtts C, Padian NS. A comparison of four condom-use measures in predicting pregnancy, cervical STI and HIV incidence among Zimbabwean women. *Sex Transm Infect.* 2010 Jun;**86**(3):231-5.
- 65 Pettifor AE, Levandowski BA, Macphail C, et al. A tale of two countries: rethinking sexual risk for HIV among young people in South Africa and the United States. *J Adolesc Health.* 2011 Sep;**49**(3):237-43 e1.
- 66 Sales JM, Brown JL, Vissman AT, DiClemente RJ. The association between alcohol use and sexual risk behaviors among African American women across three developmental periods: a review. *Curr Drug Abuse Rev.* 2012 Jun;**5**(2):117-28.
- 67 Li Q, Li X, Stanton B. Alcohol use and sexual risk behaviors and outcomes in China: a literature review. *AIDS Behav.* 2010 Dec;**14**(6):1227-36.
- 68 Rehm J, Shield KD, Joharchi N, Shuper PA. Alcohol consumption and the intention to engage in unprotected sex: systematic review and meta-analysis of experimental studies. *Addiction.* 2012 Jan;**107**(1):51-9.
- 69 Fisher JC, Bang H, Kapiga SH. The association between HIV infection and alcohol use: a systematic review and meta-analysis of African studies. *Sex Transm Dis.* 2007 Nov;**34**(11):856-63.
- 70 Cook RL, Clark DB. Is there an association between alcohol consumption and sexually transmitted diseases? A systematic review. *Sexually Transmitted Diseases.* 2005;**32**(3):156-64.
- 71 Kalichman SC, Simbayi LC, Kaufman M, Cain D, Jooste S. Alcohol use and sexual risks for HIV/AIDS in sub-Saharan Africa: systematic review of empirical findings. *Prev Sci.* 2007 Jun;**8**(2):141-51.
- 72 Pithey A, Parry C. Descriptive systematic review of sub-Saharan African studies on the association between alcohol use and HIV infection. *SAHARA J (Journal of Social Aspects of HIV/AIDS Research Alliance).* 2009;**6**(4).
- 73 Mataure P, McFarland W, Fritz K, et al. Alcohol use and high-risk sexual behavior among adolescents and young adults in Harare, Zimbabwe. *AIDS and Behavior.* 2002;**6**(3):211-9.
- 74 Chersich MF, Rees HV. Vulnerability of women in southern Africa to infection with HIV: biological determinants and priority health sector interventions. *AIDS.* 2008 Dec;**22 Suppl 4**:S27-40.
- 75 Morojele NK, Kachieng'a MA, Nkoko MA, et al. Perceived effects of alcohol use on sexual encounters among adults in South Africa. *African Journal of Drug and Alcohol Studies.* 2004;**3**:1-20.
- 76 Kalichman SC, Grebler T. Reducing numbers of sex partners: do we really need special interventions for sexual concurrency? *AIDS Behav.* 2010 Oct;**14**(5):987-90.
- 77 Pihl RO, Peterson JB, Lau MA. A biosocial model of the alcohol-aggression relationship. *Journal of Studies on Alcohol and Drugs.* 1993(**11**):128.
- 78 Steele CM, Josephs RA. Alcohol myopia: Its prized and dangerous effects. *American Psychologist.* 1990;**45**(8):921.
- 79 Graham K, West P, Wells S. Evaluating theories of alcohol-related aggression using observations of young adults in bars. *Addiction.* 2000;**95**(6):847-63.
- 80 Parry CD. South Africa: alcohol today. *Addiction.* 2005;**100**(4):426-9.
- 81 Parry CD, Myers B, Morojele NK, et al. Trends in adolescent alcohol and other drug use: findings from three sentinel sites in South Africa (1997–2001). *Journal of Adolescence.* 2004;**27**(4):429-40.
- 82 Rehm J, rgen u, Rehn N, et al. The global distribution of average volume of alcohol consumption and patterns of drinking. *European addiction research.* 2003;**9**(4):147-56.
- 83 Morojele NK, Kachieng'a MA, Mokoko E, et al. Alcohol use and sexual behaviour among risky drinkers and bar and shebeen patrons in Gauteng province, South Africa. *Social science & medicine.* 2006;**62**(1):217-27.
- 84 Cain D, Pare V, Kalichman SC, et al. HIV risks associated with patronizing alcohol serving establishments in South African Townships, Cape Town. *Prev Sci.* 2012 Dec;**13**(6):627-34.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

- 85 Weir SS, Pailman C, Mahlalela X, Coetzee N, Meidany F, Boerma JT. From people to places: focusing AIDS prevention efforts where it matters most. *AIDS*. 2003 Apr 11;**17**(6):895-903.
- 86 Wohl DA, Khan MR, Tisdale C, et al. Locating the places people meet new sexual partners in a southern US city to inform HIV/STI prevention and testing efforts. *AIDS Behav*. 2011 Feb;**15**(2):283-91.
- 87 Weir SS, Merli MG, Li J, et al. A comparison of respondent-driven and venue-based sampling of female sex workers in Liuzhou, China. *Sex Transm Infect*. 2012 Dec;**88 Suppl 2**:i95-101.
- 88 Singh K, Sambisa W, Munyati S, et al. Targeting HIV interventions for adolescent girls and young women in Southern Africa: use of the PLACE methodology in Hwange District, Zimbabwe. *AIDS Behav*. 2010 Feb;**14**(1):200-8.
- 89 Speizer IS, Beauvais H, Gomez AM, Outlaw TF, Roussel B. Using multiple sampling approaches to measure sexual risk-taking among young people in Haiti: programmatic implications. *Stud Fam Plann*. 2009 Dec;**40**(4):277-88.
- 90 Tate J, Singh K, Ndubani P, Kamwanga J, Buckner B. Measurement of HIV prevention indicators: a comparison of the PLACE method and a household survey in Zambia. *AIDS Behav*. 2010 Feb;**14**(1):209-17.
- 91 Khan MR, Rasolofomanana JR, McClamroch KJ, et al. High-risk sexual behavior at social venues in Madagascar. *Sex Transm Dis*. 2008 Aug;**35**(8):738-45.
- 92 Mwandagaliwa K, Jackson EF, McClamroch K, Bollinger R, Ryder RW, Weir SS. Local differences in human immunodeficiency virus prevalence: a comparison of social venue patrons, antenatal patients, and sexually transmitted infection patients in eastern kinshasa. *Sex Transm Dis*. 2009 Jul;**36**(7):406-12.
- 93 Figueroa JP, Weir SS, Byfield L, Hall A, Cummings SM, Suchindran CM. The challenge of promoting safe sex at sites where persons meet new sex partners in Jamaica: results of the Kingston PLACE randomized controlled trial. *Trop Med Int Health*. 2010 Aug;**15**(8):945-54.
- 94 Weir SS, Tate JE, Zhusupov B, Boerma JT. Where the action is: monitoring local trends in sexual behaviour. *Sex Transm Infect*. 2004 Dec;**80 Suppl 2**:ii63-8.
- 95 Weir SS, Morroni C, Coetzee N, Spencer J, Boerma JT. A pilot study of a rapid assessment method to identify places for AIDS prevention in Cape Town, South Africa. *Sex Transm Infect*. 2002 Apr;**78 Suppl 1**:i106-13.
- 96 Sivaram S, Srikrishnan AK, Latkin C, et al. Male alcohol use and unprotected sex with non-regular partners: evidence from wine shops in Chennai, India. *Drug Alcohol Depend*. 2008 Apr 1;**94**(1-3):133-41.
- 97 Parks KA, Hsieh YP, Collins RL, Levonyan-Radloff K, King LP. Predictors of risky sexual behavior with new and regular partners in a sample of women bar drinkers. *J Stud Alcohol Drugs*. 2009 Mar;**70**(2):197-205.
- 98 Kalichman SC, Simbayi LC, Vermaak R, Jooste S, Cain D. HIV/AIDS risks among men and women who drink at informal alcohol serving establishments (Shebeens) in Cape Town, South Africa. *Prev Sci*. 2008 Mar;**9**(1):55-62.
- 99 Fritz KE, Woelk GB, Bassett MT, et al. The association between alcohol use, sexual risk behavior, and HIV infection among men attending beerhalls in Harare, Zimbabwe. *AIDS and Behavior*. 2002;**6**(3):221-8.
- 100 Bassett MT, McFarland WC, Ray S, et al. Risk factors for HIV infection at enrollment in an urban male factory cohort in Harare, Zimbabwe. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 1996;**13**(3):287-93.
- 101 Lewis JJ, Garnett GP, Mhlanga S, Nyamukapa CA, Donnelly CA, Gregson S. Beer halls as a focus for HIV prevention activities in rural Zimbabwe. *Sex Transm Dis*. 2005 Jun;**32**(6):364-9.
- 102 Go VF, Solomon S, Srikrishnan AK, et al. HIV rates and risk behaviors are low in the general population of men in Southern India but high in alcohol venues: results from 2 probability surveys. *J Acquir Immune Defic Syndr*. 2007 Dec 1;**46**(4):491-7.
- 103 Kalichman SC, Simbayi LC, Vermaak R, et al. Randomized trial of a community-based alcohol-related HIV risk-reduction intervention for men and women in Cape Town South Africa. *Ann Behav Med*. 2008 Dec;**36**(3):270-9.
- 104 Fritz K, McFarland W, Wyrod R, et al. Evaluation of a peer network-based sexual risk reduction intervention for men in beer halls in Zimbabwe: results from a randomized controlled trial. *AIDS Behav*. 2011 Nov;**15**(8):1732-44.
- 105 Power R, Langhaug L, Cowan F. "But there are no snakes in the wood": risk mapping as an outcome measure in evaluating complex interventions. *Sex Transm Infect*. 2007 Jun;**83**(3):232-6.
- 106 Kandala NB, Ji C, Cappuccio PF, Stones RW. The epidemiology of HIV infection in Zambia. *AIDS Care*. 2008 Aug;**20**(7):812-9.
- 107 Kleinschmidt I, Pettifor A, Morris N, MacPhail C, Rees H. Geographic distribution of human immunodeficiency virus in South Africa. *Am J Trop Med Hyg*. 2007 Dec;**77**(6):1163-9.
- 108 Messina JP, Emch M, Muwonga J, et al. Spatial and socio-behavioral patterns of HIV prevalence in the Democratic Republic of Congo. *Soc Sci Med*. 2010 Oct;**71**(8):1428-35.
- 109 Moran D, Jordaan JA. HIV/AIDS in Russia: determinants of regional prevalence. *Int J Health Geogr*. 2007;**6**:22.

Dissertation Proposal: Molly Rosenberg

School Enrollment, Alcohol Outlets, and Sexual Risk

- 110 Tanser F, Lesueur D, Solarsh G, Wilkinson D. HIV heterogeneity and proximity of homestead to roads in rural South Africa: an exploration using a geographical information system. *Trop Med Int Health*. 2000 Jan;**5**(1):40-6.
- 111 Barnighausen T, Tanser F, Gqwede Z, Mbizana C, Herbst K, Newell ML. High HIV incidence in a community with high HIV prevalence in rural South Africa: findings from a prospective population-based study. *AIDS*. 2008 Jan **22**(1):139-44.
- 112 Gabrysch S, Edwards T, Glynn JR. The role of context: neighbourhood characteristics strongly influence HIV risk in young women in Ndola, Zambia. *Trop Med Int Health*. 2008 Feb;**13**(2):162-70.
- 113 Moise I, Kalipeni E. Applications of geospatial analysis to surveillance data: a spatial examination of HIV/AIDS prevalence in Zambia. *GeoJournal*. 2012 2012/08/01;**77**(4):525-40.
- 114 Campbell CA, Hahn RA, Elder R, et al. The effectiveness of limiting alcohol outlet density as a means of reducing excessive alcohol consumption and alcohol-related harms. *Am J Prev Med*. 2009 Dec;**37**(6):556-69.
- 115 Popova S, Giesbrecht N, Bekmuradov D, Patra J. Hours and days of sale and density of alcohol outlets: impacts on alcohol consumption and damage: a systematic review. *Alcohol Alcohol*. 2009 Sep-Oct;**44**(5):500-16.
- 116 Livingston M, Chikritzhs T, Room R. Changing the density of alcohol outlets to reduce alcohol-related problems. *Drug Alcohol Rev*. 2007 Sep;**26**(5):557-66.
- 117 Bryden A, Roberts B, McKee M, Petticrew M. A systematic review of the influence on alcohol use of community level availability and marketing of alcohol. *Health Place*. 2012 Mar;**18**(2):349-57.
- 118 Grube JW, Stewart K. Preventing impaired driving using alcohol policy. *Traffic Inj Prev*. 2004 Sep;**5**(3):199-207.
- 119 Alaniz ML. Alcohol availability and targeted advertising in racial/ethnic minority communities. *Alcohol Health Res World*. 1998;**22**(4):286-9.
- 120 Cook PJ, Moore MJ. The economics of alcohol abuse and alcohol-control policies. *Health Aff (Millwood)*. 2002 Mar-Apr;**21**(2):120-33.
- 121 Babor TF, Aguirre-Molina M, Marlatt GA, Clayton R. Managing alcohol problems and risky drinking. *Am J Health Promot*. 1999 Nov-Dec;**14**(2):98-103.
- 122 Room R, Babor T, Rehm J. Alcohol and public health. *Lancet*. 2005 Feb 5-11;**365**(9458):519-30.
- 123 Cohen DA, Ghosh-Dastidar B, Scribner R, et al. Alcohol outlets, gonorrhea, and the Los Angeles civil unrest: a longitudinal analysis. *Soc Sci Med*. 2006 Jun;**62**(12):3062-71.
- 124 Scribner RA, Cohen DA, Farley TA. A geographic relation between alcohol availability and gonorrhea rates. *Sex Transm Dis*. 1998 Nov;**25**(10):544-8.
- 125 Theall KP, Scribner R, Cohen D, et al. The neighborhood alcohol environment and alcohol-related morbidity. *Alcohol Alcohol*. 2009 Sep-Oct;**44**(5):491-9.
- 126 Nichols BE, Nkalamo D, Whitcomb BW. Density of drinking establishments and HIV prevalence in a migrant town in Namibia. *AIDS Behav*. 2012 May;**16**(4):829-34.
- 127 Eaton LA, Kalichman SC, Pitpitan EV, et al. The relationship between attending alcohol serving venues nearby versus distant to one's residence and sexual risk taking in a South African township. *J Behav Med*. 2013 Feb 13.
- 128 Kalichman S. Social and Structural HIV Prevention in Alcohol-Serving Establishments. *Alcohol research & health: the journal of the National Institute on Alcohol Abuse and Alcoholism*. 2009;**33**(3):184-94.
- 129 Kahn K, Collinson MA, Gomez-Olive FX, et al. Profile: Agincourt Health and Socio-demographic Surveillance System. *Int J Epidemiol*. 2012 Aug;**41**(4):988-1001.
- 130 Kahn K, Tollman SM, Collinson MA, et al. Research into health, population and social transitions in rural South Africa: data and methods of the Agincourt Health and Demographic Surveillance System. *Scand J Public Health Suppl*. 2007 Aug;**69**:8-20.
- 131 Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*. 1992;**34**(1):1-14.
- 132 Subramanian S, Jones K, Duncan C. Multilevel methods for public health research. In: Kawachi I, Berkman L, editors. *Neighborhoods and Health*. New York, NY: Oxford University Press; 2003.
- 133 Diez Roux AV, Aiello AE. Multilevel analysis of infectious diseases. *J Infect Dis*. 2005 Feb 1;**191** Suppl 1:S25-33.
- 134 Todd J, Carpenter L, Li X, Nakiyingi J, Gray R, Hayes R. The effects of alternative study designs on the power of community randomized trials: evidence from three studies of human immunodeficiency virus prevention in East Africa. *Int J Epidemiol*. 2003 Oct;**32**(5):755-62.
- 135 Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med*. 2004 May-Jun;**2**(3):204-8.

Dissertation Proposal: Molly Rosenberg
School Enrollment, Alcohol Outlets, and Sexual Risk

136 Morgan D, Mahe C, Mayanja B, Okongo JM, Lubega R, Whitworth JA. HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? AIDS. 2002 Mar 8;**16**(4):597-603.

137 Raguin G, Malkin JE. Genital herpes: epidemiology and pathophysiology. Update and new perspectives. Ann Med Interne (Paris). 1997;**148**(8):530-3.