**WEB ONLY APPENDIX: Relationships between Neighbourhood Characteristics and Current STI Status among HIV-Infected and HIV-Uninfected Women Living in the Southern United States: A Cross-Sectional Multilevel Analysis**

**METHODS**

**Sample and Recruitment**

The Women’s Interagency HIV Study (WIHS) is a multisite, prospective cohort study designed to characterize the impact and progression of HIV among US women.22 In 2013, WIHS expanded to clinical research sites in the Southern US (i.e., Alabama, Florida, Georgia, Mississippi, and North Carolina). These sites enrolled HIV-infected women and HIV-uninfected women at high risk of HIV acquisition between October 2013 and September 2015. WIHS eligibility criteria included being a woman between 25-60 years old. In addition, HIV-infected women were antiretroviral therapy (ART) naïve or started highly active antiretroviral therapy (HAART) after December 31, 2004; had never used didanosine, zalcitabine, or stavudine (unless during pregnancy or for pre- or post-exposure HIV prophylaxis); had never been on non-HAART ART, and had documented pre-HAART CD4 counts and HIV viral load. HIV-uninfected women reported either she or her sexual partner met at least one of the following criteria in the last five years: clinical STI diagnosis; using injection drugs, crack, cocaine, heroin, or methamphetamine; sex for drugs, money, or shelter; unprotected sex with three or more partners; sex with six or more partners; or sex with a known HIV-positive partner.

Participants were recruited by WIHS using several strategies, including clinic and community-based organization referrals. Institutional Review Board (IRB) approval was obtained at each of the collaborating institutions and written informed consent was obtained prior to initiation of study procedures. Women were compensated for time and travel. Methods are described in detail elsewhere.22 This secondary analysis, approved by the Emory University IRB, is restricted to WIHS participants who provided written informed consent to collect and geocode their residential address.

**Data Collection and Measures**

Outcome: Current Sexually Transmitted Infection

The outcome, testing positive for a current STI (binary), was defined as a laboratory-confirmed diagnosis for at least one of the following at baseline: chlamydia, gonorrhoea, trichomoniasis, or early syphilis (titre and confirmatory test results consistent with primary, secondary, or early latent [<1 year duration] infection). Assessment for each STI was conducted according to WIHS protocol-requirements (Table 1). Participants with a current STI were referred to medical providers for treatment.

**Table 1. Specimen type, test, and sensitivity and specificity for sexually transmitted infection (STI) evaluation at each Women’s Interagency HIV Study Southern Site**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **STI** | **Specimen Type** | **Site** | | | | |
| **Alabama** | **Florida** | **Georgia** | **North Carolina** | **Mississippi** |
| **Test (Sensitivity, Specificity)** | | | | |
| Chlamydia | Cervical swab1 | APTIMA Nucleic Acid Amplification Test (NAAT) (94.30, 98.00) | Becton Dickinson (BD) Probetec ET System (93.80, 99.80) | Aptima Combo 2 for CT/NG (96.60, 98.50) | Gen Probe Aptima (98.30, 96.10) | ROCHE Cobas polymerase chain reaction (PCR) (94.90, 99.40) |
| Gonorrhoea | Cervical swab1 | APTIMA NAAT (92.00, 99.80) | BD Probetec ET System (88.00, 99.80) | Aptima Combo 2 for CT/NG (96.60, 98.50) | Gen Probe Aptima (97.30, 99.00) | ROCHE Cobas PCR (96.60, 99.90 |
| Trichomoniasis | Vaginal swab2 | APTIMA NAAT (100.00, 98.10) | Wet mount3 (N/A) | Wet mount3 (N/A) | Wet mount3 (N/A) | Wet mount3 (N/A) |
| Early Syphilis4 | Serum | BD Screening rapid plasma regain (RPR) with confirmatory Treponema pallidum haemagglutination assay or Treponema pallidum particle agglutination assay (Screening: 3.00-100.00, 98.00, Confirmatory: 99.40, 100.00) | Arlington Scientific RPR Card (95.00, 98.00) | BD RPR titre, with confirmatory IgG enzyme immunoassay if reactive (N/A) | Labcorp Screening RPR with Confirmatory Quantitative RPR (Screening: 99.00, 98.40; Confirmatory: 100.00, 99.80) | BD RPR (86.00, N/A) |

1Urine was used for testing at Alabama site.

2Cervical swab was used for testing at Alabama site

3Trichomoniasis is defined as the presence of motile trichomonads on a vaginal wet mount.

4Titer and confirmatory testing results consistent with primary, secondary, or early latent (<1 year duration) infection.

Because the WIHS protocol specified collection of cervical swabs for *Chlamydia trachomatis* and *Neisseria gonorrhoeae* testing, women were not tested for these pathogens if they had undergone a hysterectomy; we classified these test results as missing. Participants who did not have laboratory test results for all four of the STIs available and also did not have at least one positive test among available results were excluded from the analyses (n=63, 8.55%). Study sites other than Georgia had higher rates of missing STI data, as did HIV-infected participants (p<0.05). We used maximum likelihood estimation and included covariates (i.e., HIV status, enrolment site) associated with missing outcome data in all multivariable models, thus controlling for potential bias introduced.23

Census Tract-Level Characteristics

Baseline participant residential addresses were geocoded to census tracts. Measures describing the social and physical environments of the census tracts where women lived were constructed using existing data sources (e.g., US Census) as follows (Table 2):

The 2013 American Community Survey (ACS) 5-year tract estimates were used to determine the percentage of residents living in poverty, percentage of unemployed residents, and percentage renter-occupied housing units.

The percentage of vacant housing units was obtained from the Vacant Address Database, a collaboration by the US Postal Services and the US Department of Housing and Urban Development.

The total number of newly reported cases of primary and secondary syphilis, gonorrhoea, and chlamydia per census tract in 2013 was obtained from State Health Departments and was used to calculate the STI prevalence per 1,000 residents aged 15-64 for each tract in 2013.

Alcohol outlet density (i.e., the number of businesses with a license to sell alcohol for off-premises per square mile in 2014) was created by geocoding address data obtained from state licensing agencies to census tract boundaries. The locations of violent crimes (i.e., murder, non-negligent manslaughter, forcible rape, robbery, and aggravated assault) during 2013 were obtained from law enforcement agencies, geocoded to census tracts, and used to calculate the violent crime rate per 1000 residents for each tract. For alcohol outlet density and violent crime rate; outlets/offenses within a 100-foot buffer of the tract boundary were included in the tract’s calculation.

**Table 2. Census tract measures, definition, data source, and year**

|  |  |  |  |
| --- | --- | --- | --- |
| **Measure** | **Definition** | **Data Source** | **Year** |
| ***Social disorder component*** | | | |
| Percent vacant housing units | Percent vacant residential housing units | United States (US) Department of Housing and Urban Development and US Postal Service | 2013 |
| Violent crime rate | Total murder, non-negligent manslaughter, forcible rape, robbery, and aggravated assaults per 1,000 tract residents1 | Law Enforcement Agencies (i.e., police department, Sheriff’s Office) | 2013 |
| Sexually transmitted infection (STI) prevalence | Prevalence of newly reported STIs (i.e., chlamydia, gonorrhoea, and primary and secondary syphilis) per 1,000 tract residents aged 15-642 | State Department of Health | 2013 |
| Percent poverty | Percent residents with annual income below poverty level | American Community Survey (ACS) | 2009-2013 |
| Percent unemployment | Percent unemployed residents ≥ 16 years old | ACS | 2009-2013 |
| ***Social disadvantage component*** | | | |
| Percent renter-occupied housing units | Percent renter occupied housing units | ACS | 2009-2013 |
| Alcohol outlet density | The number of businesses with a license to sell beverages containing alcohol (e.g., liquor, beer, wine) for off-premise consumption per tract square mile1,3 | State Licensing Agencies (e.g., Department of Revenue, Alcoholic Beverage Control Commission) | 2014 |

1Addresses were obtained from state agencies and geocoded to tracts; addresses within a 100-foot buffer of the tract boundary were included in the tract’s calculation.

2In Alabama, the number of newly identified STIs was available by ZIP code, but not census tract. ZIP-level STI counts were allocated to tracts based on the proportion of residential population using the 2015 boundaries USPS-HUD ZIP to tract crosswalk file. Twelve ZIP code-census tract combinations were not included in the crosswalk file. For these 15 participants (17% of participants with available census tract data at site), ZIP code STI prevalence was assigned to the participant census tract. We conducted sensitivity analyses, removing these participants from the analytic data set, to explore potential bias introduced by this substitution. The rounded Full Model odds ratio estimate with and without these 15 participants was the same.

3In Mississippi, off-premise liquor licensing data were available (liquor can only be purchased at package/liquor stores), but licensing data for sale of beer and wine off-premise were not publically available. As a proxy, we used non-restaurant businesses with permits to sell eggs or milk (e.g., convenience stores, pharmacies) under the oversight of the Mississippi Department of Agriculture and Commerce because these types of businesses would have refrigerated display cases and likely have the capacity to sell beer and wine.

Because several tract-level predictors were correlated, we were unable to group tract-level predictors a priori. Instead, we used principal components analysis with orthogonal rotation (varimax) to condense tract-level variables into components and to avoid multicollinearity in multivariable models. We extracted two continuous standardized principal component scores for factors with eigenvalues >1.0: (1) “social disorder” (vacant housing units, violent crime rate, STI prevalence, percent poverty, percent unemployed); and (2) “social disadvantage” (percent renter-occupied housing units, alcohol outlet density). For each factor, a one standard deviation increase indicates greater than average social disorder or social disadvantage for the sample.

Participant-Level Characteristics

WIHS collected all demographic and behavioural data using interviewer-administered questionnaires. Participant-level characteristics that might confound or modify relationships between tract-level characteristics and current STI status were determined a priori via a literature review.3 4 18

The effect modifier of interest was being HIV-infected, defined by WIHS as a reactive serologic enzyme-linked immunosorbent assay test and a confirmed positive western blot or detectable plasma HIV-1 ribonucleic acid.

Control variables captured demographic characteristics and behaviours in the past six months and were binary unless otherwise noted: age in years (continuous mean-centred); married or cohabitating; non-Hispanic African-American; annual income ≤$18,000; self-rated quality of life (QOL) (continuous mean-centred, measured using an abbreviated Medical Outcomes Study Scale ranging from 0-100, with higher scores indicative of greater QOL24); alcohol and substance use (>7 drinks in the past week or any injection or non-injection use of crack, cocaine, heroin, marijuana, hallucinogens, club drugs, methamphetamines, or recreational prescription drug use in the last 6 months); exchange of sex for drugs, money or housing; homeless (currently living in a rooming or halfway house, shelter, welfare hotel, or on the street); lifetime history of chlamydia, gonorrhoea, trichomoniasis, or syphilis; condomless vaginal or anal sex; and study site (5-level categorical).

**Analysis**

We used descriptive statistics to describe distributions of participant- and census tract-level variables at baseline. We modelled bivariate and multivariable relationships with hierarchical generalized linear models (HGLMs) using a logit link with random effects for the intercept in order to model participant-level clustering within census tracts.25 All HGLMs had two levels: participants (Level 1) were nested in census tracts (Level 2). Modelling was conducted in four stages.

In Stage 1, we used an unconditional model to assess the proportion of variance in the odds of testing positive for an STI attributable to clustering within census tracts.

In Stage 2, we modelled bivariate relationships between each tract- and participant-level characteristic and testing positive for an STI.

In Stage 3, we modelled multivariable relationships between tract-level characteristics and STIs, controlling for potential participant-level confounders. Because an aim of this study was to determine whether the magnitudes and directions of relationships between tract characteristics and having a current STI might vary by HIV status, we tested statistically for multiplicative and additive interactions between neighbourhood characteristics and testing positive for an STI by HIV status. In Stage 3A, we tested for interaction between tract characteristics and HIV status on the multiplicative scale by entering cross-level interaction terms for HIV status and tract-level variables (i.e., HIV status\*social disorder, HIV status\*social disadvantage), retaining interaction terms with p<0.05 in the multivariable model. In Stage 3B, we tested for interaction between neighbourhood characteristics and HIV status on the additive scale by fitting separate models using a binomial distribution and identity link, controlling for participant-level confounders.26 We entered cross-level interaction terms for HIV status and tract-level variables (i.e., HIV status\*social disorder, HIV status\*social disadvantage) stepwise. Interaction terms with p<0.05 were considered statistically significant on the additive scale.

Participant-level covariates traditionally included in models evaluating STI outcomes (e.g., alcohol and substance use) may lie in the causal pathways between tract characteristics and current STI status.15 Including these variables in the Full Model would attenuate relationships between tract characteristics and outcomes if they did indeed lie on the causal pathway. In Stage 4 (Reduced Model), we therefore re-ran the Full Multivariable Model excluding variables that might lie on the causal pathway between neighbourhood characteristics and STIs (i.e., income, QOL, alcohol and substance use, homelessness, STI history, unprotected sex). We compared odds ratio (OR) estimates for all tract-level variables in the Full vs. Reduced Models; differences in magnitude >10% suggested that excluded variables may attenuate relationships between neighbourhood characteristics and STIs.

HGLMs were fit using PROC GLIMMIX using Newton Raphson optimization and Gauss-Hermite quadrature approximation in SAS 9.4.

**Power Calculations**

We conducted a priori power calculations, running 1,000 repetitions for a range of binary outcome distributions that are supported by the literature, using an expected sample size of 800 and simulation methods.4 We assessed power for a variety of average census tract cluster sizes ranging from 5 to 40. Given these conditions, we had power to detect small individual-level effects (d = 0.20), small census tract-level effects (d = 0.25), as well as small cross-level interaction effects (d = 0.20) in cross-sectional analyses either on the individual- or cross-level.

**RESULTS**

**Characteristics of census tracts and participants**

A total of 845 women were enrolled at WIHS’s southern sites. One hundred eight women were excluded from these analyses because they did not have geocoded address data; the majority of these women did not consent for geocoding (n=65, 60.2%). Participants excluded from these analyses because they were missing geocoded address data were more likely to report annual household incomes ≤$18,000 (83.2% vs. 69.0%, p=0.003); alcohol and substance use (48.1% vs. 37.9%, p=0.04); and sex exchange (17.6% vs. 5.7%, p<0.0001).

In the final analytic sample (N=737), participants were on average 44 years old (SD=9.3), 71.9% were HIV-infected, and 83.3% identified as non-Hispanic African American (Table 3).

**Table 3. Distributions of census tract and participant characteristics among 737 women enrolled in the Women’s Interagency HIV Study Southern Sites**

|  |  |
| --- | --- |
| **Characteristics of participants and census tracts** | **n (%) or Mean (SD)** |
| ***Outcomes*** | |
| Laboratory confirmed STI  HIV-infected  HIV-uninfected  Overall  Missing | 58 (10.9)1  21 (10.1)  79 (10.7)  63 (8.5) |
| Chlamydia  HIV-infected  HIV-uninfected  Overall  Missing | 5 (0.9)  4 (1.9)  9 (1.2)  32 (4.3) |
| Gonorrhoea  HIV-infected  HIV-uninfected  Overall  Missing | 5 (0.9)  1 (0.5)  6 (0.8)  37 (5.0) |
| Trichomoniasis  HIV-infected  HIV-uninfected  Overall  Missing | 34 (6.4)  12 (5.8)  46 (6.2)  15 (2.0) |
| Syphilis  HIV-infected  HIV-uninfected  Overall  Missing | 17 (3.2)  4 (1.9)  21 (2.8)  19 (2.6) |
| ***Census tract-level characteristics*** | |
| *Social disorder component* | |
| Percent vacant housing units | 7.8 (6.3) |
| Violent crime rate per 1,000 residents  Missing | 13.5 (13.4)  44 (6.0) |
| Percent poverty | 29.1 (13.6) |
| Percent unemployed | 16.1 (8.0) |
| STI prevalence per 1,000 residents  Missing | 19.1 (13.3)  1 (0.1) |
| *Social disadvantage component* | |
| Percent renter-occupied housing units | 51.9 (21.7) |
| Alcohol outlet density  Missing | 4.8 (7.6)  1 (0.1) |
| ***Participant-level characteristics*** | |
| HIV-infected | 530 (71.9) |
| Age in years | 43.7 (9.3) |

|  |  |
| --- | --- |
| Marital status  Married or living as married  Widowed, divorced, or separated  Never married  Other  Missing | 244 (34.1)  221 (30.0)  265 (36.0)  4 (0.5)  3 (0.4) |
| Ethnicity/Race  Non-Hispanic African American  Non-Hispanic White  Hispanic White  Hispanic African American  Hispanic Other  Asian Pacific Islander  Native American/Alaskan  Other | 614 (83.3)  67 (9.1)  17 (2.3)  12 (1.6)  13 (1.8)  1 (0.1)  11 (1.5)  2 (0.27) |
| Annual household income  $18,000 or less  $18,001-24,000  $24,001-36,000  $36,001-75,000  >$75,000  Missing | 492 (66.8)  85 (11.5)  80 (10.8)  50 (6.8)  6 (0.8)  24 (3.3) |
| Quality of life index  Missing | 67.1 (20.5)  3 (0.4) |
| Alcohol or illicit substance use  Missing | 279 (37.9)  1 (0.2) |
| Sex exchange  Missing | 42 (5.7)  1 (0.1) |
| Homeless  Missing | 47 (6.4)  12 (1.6) |
| Lifetime STI diagnosis | 472 (64.0) |
| Condomless vaginal or anal sex  Missing | 274 (42.3)  4 (0.5) |

1Participants could test positive for more than one STI. As a result, the total count of STIs by type exceeds the total count of participants with a diagnosed STI.

Eleven percent of participants tested positive for at least one STI at baseline: 1.2% tested positive for CT, 0.8% tested positive for gonorrhoea, 6.2% tested positive for TV, and 2.8% tested positive for syphilis. The proportion testing positive for any STI was comparable by HIV status. Participants on average lived in census tracts with 29.1% of residents living in poverty (SD=13.6), and with 19.1 newly reported STI cases per 1,000 residents (SD=13.3) and 13.7 violent crimes per 1,000 residents annually (SD= 13.4) (Table 3). On average, tracts contained roughly five alcohol outlets per square mile (SD=7.8).

**Associations between neighbourhood characteristics and current STI status**

Clustering by census tracts accounted for 24.0% of the variance in the odds of having a current STI (unconditional model random intercept variance=1.04, p=0.08). In bivariate analyses (Table 4), neither tract-level social disorder (OR=1.18, 95% Confidence Interval [CI]=0.91-1.52) nor social disadvantage (OR=1.21, 95% CI=0.94-1.54) were associated with current STI status.

**Table 4. Bivariate and multivariable relationships between census tract characteristics and the odds of having a current STI among women enrolled in the Women’s Interagency HIV Study Southern Sites**

|  |  |  |  |
| --- | --- | --- | --- |
| **Characteristics of census tracts and participants** | **Bivariate**  **OR (95%)** | **Reduced Model**  **aOR (95%)1** | **Full Model**  **aOR (95%)1** |
| ***Census tract-level characteristics*** | | | |
| Social disorder component2 | 1.18 (0.91, 1.52) | **1.30 (0.99, 1.72)** | 1.25 (0.94, 1.66) |
| Social disadvantage component3 | 1.21 (0.94, 1.54) | **1.34 (0.96, 1.87)** | **1.34 (0.96, 1.86)** |
| ***Participant-level characteristics*** | | | |
| HIV-infected | 1.22 (0.68, 2.17) | 1.34 (0.71, 2.51) | 1.51 (0.76, 3.00) |
| Age in years | 1.03 (1.00, 1.06) | 1.02 (0.99, 1.05) | 1.02 (0.99, 1.05) |
| Married or living as married | 0.76 (0.44, 1.33) | 0.99 (0.55, 1.79) | 0.95 (0.52, 1.74) |
| Non-Hispanic African-American | 1.51 (0.69, 3.31) | 0.88 (0.38, 2.02) | 0.87 (0.38, 1.99) |
| Annual household income of $18,000 or less | 1.24 (0.70, 2.19) | -- | 0.86 (0.46, 1.58) |
| Quality of Life Index | 1.00 (0.99, 1.01) | -- | 1.00 (0.99, 1.02) |
| Alcohol or illicit substance use | 1.56 (0.93, 2.63) | -- | 1.46 (0.82, 2.60) |
| Sex exchange | 1.05 (0.36, 3.08) | -- | 1.12 (0.33, 3.79) |
| Homeless | 1.10 (0.40, 2.94) | -- | 1.24 (0.43, 3.46) |
| Lifetime STI diagnosis | 1.12 (0.65, 1.92) | -- | 1.13 (0.61, 2.10) |
| Condomless vaginal or anal sex | 0.92 (0.54, 1.56) | -- | 1.19 (0.65, 2.20) |
| Study site (ref=Georgia)  Alabama  Florida  Mississippi  North Carolina | 0.28 (0.10, 0.77)  0.54 (0.24, 1.21)  0.55 (0.27, 1.16)  0.22 (0.10, 0.50) | 0.17 (0.05, 0.61)  0.30 (0.10, 0.91)  0.60 (0.27, 1.33)  0.24 (0.10, 0.59) | 0.17 (0.05, 0.60)  0.32 (0.10, 0.99)  0.64 (0.29, 1.44)  0.24 (0.10, 0.59) |
| ***Model fit*** | | | |
| Random intercept variance (p-value) | -- | 0.39 (0.24) | 0.32 (0.29) |
| -2LL | -- | 403.41 | 400.28 |
| AIC |  | 427.41 | 438.28 |
| BIC |  | 475.27 | 514.07 |

1Multivariable modelling was restricted to participants with no missing data for predictor or outcome variables (n=589).

2Component generated from principal component analysis (PCA) including tract-level vacant housing units, violent crime rate, STI prevalence, poverty, and unemployment.

3Component generated from PCA including tract-level renter-occupied housing and alcohol outlet density.

In the Reduced Model, tract-level social disorder (OR=1.30, 95% CI=0.99-1.72, p=0.06) was associated with having a current STI when excluding participant-level covariates that may lie on the causal pathway. Tract-level social disadvantage (OR=1.34, 95% CI=0.96-1.87, p=0.08) was associated with having a current STI (Reduced Model). Specifically, a one standard deviation higher social disorder component was associated with a 30% greater odds of having a current STI and a one standard deviation higher social disadvantage component was associated with a 34% greater odds of having a current STI.

There was no evidence of effect modification of the relationships between social disorder, social disadvantage, and having a current STI by HIV status on either a multiplicative or additive scale (p>0.05). Odds ratios for tract-level characteristics in the Full Model, as compared to the Reduced Model excluding participant-level variables that might lie on the causal pathway were within 4% for all comparisons, suggesting that excluded variables did not lie on the causal pathway.