Beyond the Individual: The Roles of Social and Structural Contexts in HIV Prevention and HIV Acquisition in the United States

Jessica L. Maksut

University of Connecticut - Storrs, jessica.maksut@uconn.edu

Follow this and additional works at: https://opencommons.uconn.edu/dissertations

Recommended Citation
https://opencommons.uconn.edu/dissertations/1876
Beyond the Individual: The Roles of Social and Structural Contexts in HIV Prevention and HIV Acquisition in the United States

Jessica L. Maksut, Ph.D.

University of Connecticut, 2018

The present dissertation leverages the utility of analyzing a variety of secondary data sources to explore relationships between qualities of social and structural environments and HIV-related outcomes via a social-ecological approach. Advances in HIV outcomes for vulnerable, high-risk populations, e.g., African-American/Black gay, bisexual, and other men who have sex with men (BMSM), have been stymied by social and behavioral scientists’ tendency to primarily call upon individual-behavioral factors to explain the elevated rates of HIV observed within BMSM communities. This dissertation employs a broader analytical lens to explore relationships between social and structural variables with HIV acquisition and other key HIV-related outcomes (Studies 1 and 2).

This dissertation also uses social media data to garner insights about the general public’s understandings of, and attitudes toward, extant HIV prevention tools. Location-based social media data are used to link attitudes toward HIV prevention tools with various social and structural characteristics of the geographic locations from where the content originates (Study 3).

The results of the three studies indicate that there are real HIV prevention and acquisition considerations for social- and structural-level variables, such that factors at these levels have significant main and interactive associations with key HIV prevention, HIV risk behavior, and HIV acquisition variables. Taken together, the results indicate that HIV prevention and care strategies should not treat HIV as an independent social problem. Instead, future interventions
must be multi-level in nature, with goals of positive behavioral as well as social and structural change.
Beyond the Individual: The Roles of Social and Structural Contexts in HIV Prevention and HIV Acquisition in the United States

Jessica L. Maksut

B.A., Quinnipiac University, 2012
M.A., University of Connecticut, 2015

A Dissertation
Submitted in Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy

University of Connecticut
2018
Copyright by

Jessica Maksut
Beyond the Individual: The Roles of Social and Structural Contexts in HIV Prevention and HIV Acquisition in the United States

Presented by
Jessica L. Maksut, B.A., M.A.

Major Advisor
Lisa Eaton, Ph.D.

Associate Advisor
Edna Brown, Ph.D.

Associate Advisor
Keith Bellizzi, Ph.D.

University of Connecticut
2018
Acknowledgments

I would like to express my deepest gratitude toward a number of special individuals who are responsible for sustaining me during my time at UConn, and especially during the final, dissertation stage of my program. The overwhelming amount of support I received from family, friends, colleagues, and mentors was, in effect, the fuel I needed to accomplish my goals as a graduate student and to take the next, very exciting, step forward along my career trajectory in Baltimore.

First and foremost, I want to acknowledge my major advisor, Dr. Lisa Eaton, for whom I have the utmost respect and who gave me a quality of mentorship and guidance that is unmatched. Thank you, Lisa, for making my time at UConn meaningful, for helping me to grow as a scientist, and for being a wonderful role model. I aspire to be as generous, well-respected, and hardworking as you are, and for the work I do to be as impactful as yours has been.

Second, I want to express my gratitude toward my dissertation committee members, Dr. Edna Brown and Dr. Keith Bellizzi, who were generous with their time and who offered insightful feedback and suggestions to make this dissertation as thoughtful, thorough, and informed as possible. Additionally, I want to thank Dr. Seth Kalichman for teaching me so much of what I now know about HIV social and behavioral science, and for making my T32 experience a great one.

Third, I want to thank Daniel Snyder and Michael Tynes for offering me the benefits of their skill sets, patience, and support, particularly with the programming and statistical aspects of the Twitter project. Both of you were key to seeing the project through to its completion, and I look forward to seeing all of the great things each of you accomplishes in the future.
Fourth, I want to thank my friends and colleagues, Chanee Fabius, Liz Siembida, Kate Dibble, and Kristen DiFilippo for their constant presence, kindness, and emotional support. I consider myself very lucky to be surrounded by such special, kind, and powerful women.

Fifth, I want to acknowledge my parents and my brother, C.J., who together helped me to cultivate the person I am today. Mom, from you I learned how to be strong by being empathetic and to love without reservation. Dad, from you I learned how to listen and to always prioritize the people you love. C.J., from you I learned how to not take myself too seriously and to always share my accomplishments with others. All my love.

Finally, I owe an especially large thank you to my incredibly supportive and loving partner, Niranjan, who has, without a single complaint, kept me strong and kept me together during this particularly stressful time. I look forward to a lifetime of making you laugh until you cry.
Dedication

For all women who have changed, are changing, and will change the world with science.
# Table of Contents

Acknowledgments...........................................................................................................iv
Dedication............................................................................................................................vi
Table of Contents................................................................................................................vii
List of Tables and Figures....................................................................................................viii

Chapter 1: Introduction......................................................................................................1

Chapter 2: Proximal and underlying determinants of HIV diagnosis among Black gay, bisexual, and other men who have sex with men in Atlanta, GA........................................24

Chapter 3: Neighborhood poverty, enacted stigma from health care providers, and HIV/STI among Black gay, bisexual, and other men who have sex with men in Atlanta, Georgia..........................................................................................................................54

Chapter 4: Using Twitter for remote monitoring of public attitudes toward, and activity around, antiretroviral pre-exposure prophylaxis (PrEP) for HIV prevention.................................................................................................................................................78

Chapter 5: Discussion........................................................................................................100
List of Tables and Figures

**Introduction:**
Figure 1. Study 1’s Theoretical Framework..........................................................21
Figure 2. Study 2’s Theoretical Framework..........................................................22
Figure 3. Study 3’s Theoretical Framework..........................................................23

**Study 1:**
Table 1. HIV Risk Behaviors by HIV Status at the 12-month Follow-up Appointment........50
Table 2. Individual-level and Area-level Factors by HIV Status at the 12-month Follow-up.......51
Table 3. Fixed-effects Estimates for Models 1-4 with HIV Status as the Outcome Variable.......52

**Study 2:**
Table 1. Descriptive Statistics for the Sample of 147 Black Men Who Have Sex with Men (BMSM) Located in the Atlanta, GA Metropolitan Area...............................................................75
Table 2. Results from the Fixed Effects Regression Analyses Showing the Moderation Effect of Population Density on the Relationship Between Neighborhood Poverty and HIV Infection.........................................................................................................................76
Table 3. Results from the PAF Calculations for Enacted Stigma and HIV and STI Cases........77

**Study 3:**
Table 1. Pearson Correlations (r) Between the Independent and Dependent Variables for the Geolocated Tweets in the U.S. MSAs of Interest........................................................................................................98
Table 2. Fixed Effects Regressions Results of Factors Associated with Tweet Frequency and Tweet Sentiment.........................................................................................................................99
Chapter 1: Introduction

In 2016, 39,782 people were diagnosed with human immunodeficiency virus (HIV) in the United States (Centers for Disease Control and Prevention (CDC), 2017). While the annual number of infections decreased by five percent from 2011 to 2015, not all socio-demographic groups have experienced the same decrease. Gay, bisexual, and other men who have sex with men (MSM) continue to bear the greatest burden of all risk groups, representing an estimated 26,200 (69.3%) of these new infections. Among MSM, the most affected groups include MSM of color, particularly MSM who are African-American or Black (i.e., BMSM), who accounted for 38 percent of new infections among MSM and 26 percent of all new infections in 2016 (CDC, 2017). Even more troubling is the fact that HIV infections among BMSM who are younger appear to be on the rise. In 2017, the CDC reported that HIV diagnoses among young adult (aged 25-34) BMSM increased by 30 percent from 2011 to 2015, while during the same time period, HIV diagnoses among middle adult (45-54) BMSM decreased by 25 percent. In order to sustain and accelerate our advances toward eradicating HIV in the U.S., we must attend further to the needs of young adult BMSM who are at risk for HIV infection.

Burden of HIV Among Black Men Who Have Sex with Men

Since the beginning of the HIV/AIDS epidemic in the U.S., African-American/Black persons, in particular those who are MSM (i.e., BMSM) have been heavily, and disproportionately, burdened by the disease. Even today, the severity of the burden of HIV among BMSM communities cannot be understated. While BMSM comprise an estimated 0.2 percent of the U.S. population, they account for 22 percent of new HIV infections annually (CDC, 2010; Rosenberg, Millett, Sullivan, Del Rio, & Curran, 2014). In recent years, HIV infections among young adult BMSM have increased by approximately 30 percent (CDC, 2017) and, if current rates continue,
an estimated 60 percent of BMSM will be living with HIV by the time they reach 40 years of age (Matthews et al., 2016). Taken together, prior research demonstrates that the alarming rates of HIV infection among BMSM, especially younger BMSM, are likely to persist or even increase if we do not direct our attention and energy toward the provision of resources to communities most in need, as well as develop innovative strategies to better understand and address these epidemics.

According to the CDC (2017), racial disparities in HIV are alarmingly pronounced, with BMSM having HIV incidence and prevalence rates that are significantly higher than their counterparts of other racial/ethnic backgrounds. Fields et al. (2012) found that BMSM are five times more likely than White MSM to be living with HIV despite statistically equivalent rates of HIV risk behaviors, including sexual risk behavior (e.g., condomless anal intercourse [CAI]). In the past, researchers have generally agreed upon CAI being the “prime candidate for the disparate infection rates between BMSM and MSM of other racial/ethnic groups” (Millett et al., 2007, p. 2083). Therefore, social and behavioral HIV scientists have tended toward intervening upon individual-level risk factors (e.g., CAI, drug or alcohol use during sexual activity, sexually transmitted infections (STIs)) in an attempt to see improvements in HIV-related outcomes (e.g., HIV risk, HIV acquisition) among BMSM (Millett, Peterson, Wolitski, & Stall, 2006; Oster et al., 2011). However, focusing only on individual-level risk factors (i.e., drug- and sex-related HIV risk behaviors) ensures that we will continue to fail to (1) adequately address the needs of BMSM, and (2) make progress in eliminating the HIV epidemic in this population. Given that BMSM do not engage in HIV risk behaviors more frequently than their counterparts with lower rates of HIV infection (Fields et al., 2012; Millett et al., 2007), a broader lens – which considers factors beyond the individual-level – must be employed to detect and address the determinants of this disparity.
Indeed, despite HIV prevention scientists’ best efforts to address individual behavioral risk factors, the HIV epidemic among BMSM has persisted (Maulsby et al., 2014; Rosenberg et al., 2014). A likely explanation for these efforts falling short is that BMSM have been found to engage in HIV risk behaviors, including CAI, at comparable or lower rates than other MSM (Harawa et al., 2004; Koblin et al., 2006). However, despite the lack of evidence that individual-level factors fully account for BMSM’s elevated rates of HIV, too few social and behavioral HIV scientists have moved beyond individual-level risk factors to also consider, from a social-ecological perspective, whether and to what extent factors at other levels of influence (e.g., interpersonal, community, institutional, and/or policy levels) are associated with BMSM’s elevated HIV infection rates (Levy et al., 2014; Phillips, Birkett, Kuhns, Hatchel, Garofalo, & Mustanski, 2015; Sumartojo, 2000). While there is relatively unanimous agreement that social and structural factors play a role in BMSM’s HIV-related outcomes, including their elevated HIV incidence and prevalence rates (Sevelius, Grinstead Reznick, Hart, & Schwarcz, 2009), these factors are all too often deemed to be “outside of the purview of behavioral interventionists” (Latkin, Weeks, Glasman, Galletly, & Albarracin, 2010), and are therefore infrequently studied and addressed.

**HIV Risk: Relevant Social and Structural Factors**

The lack of attention to other social-ecological levels of influence is concerning for reasons outlined in a seminal systematic review by Millett et al. (2012). The authors found that HIV-related racial/ethnic disparities between BMSM and MSM of other races/ethnicities are greatest in the following areas: (1) BMSM have greater numbers of STI diagnoses and undiagnosed HIV infections than MSM of other races/ethnicities, (2) BMSM are, on average, less likely to be virally suppressed than MSM of other races/ethnicities, (3) BMSM have less access to antiretroviral medications (ART) than MSM of other races/ethnicities, (4) BMSM are less likely to be adherent
to ART than MSM of other races/ethnicities, and finally, (5) BMSM are less likely to attend their HIV care appointments as compared to MSM of other races/ethnicities. BMSM’s higher rates of STIs and higher viral loads, as compared to their counterparts of other races/ethnicities, is troubling for a number reasons. First, there is a wealth of evidence which suggests that STI infections and high viral loads promote HIV transmission. The World Health Organization (WHO, 2018) reported that the presence of an untreated STI can increase one’s risk of becoming infected with HIV by a factor of up to 10. Further, persons living with HIV (PLWH) are more likely to transmit the infection to their sexual partner if either of them already has an untreated STI. Second, PLWH with higher viral loads – meaning a greater amount of virus in their bodies – are more likely to transmit HIV to their partners (WHO, 2018; Wilson et al., 2008). Third, these findings together suggest that BMSM, both HIV positive and HIV negative, are likely less connected to care than their counterparts of other races/ethnicities. Higher rates of untreated STIs and higher viral loads indicate that HIV positive BMSM may have more difficulty remaining adherent to their HIV medications, and that both HIV positive and HIV negative BMSM have untreated, potentially unrecognized STIs at higher rates, perhaps due to less frequent care visits which during which time infections could be detected and treated. In sum, differences in access to, and engagement in, care may be at least partially responsible for elevated rates of HIV among BMSM.

In a study by Phillips and his colleagues (2015), for MSM residing in Chicago (not BMSM specific), walkability of one’s neighborhood (which may impact accessibility of health care resources) was strongly associated with HIV infection. Furthermore, Neaigus et al. (2013)’s study of BMSM in New York City found that participants residing in Brooklyn were most at risk for HIV infection, and they hypothesized that this phenomenon occurred at least in part because there were a high number of undiagnosed HIV infections in Brooklyn, and also due to the high
prevalence of HIV among MSM in that area. Taken together, these studies suggest that accessibility of HIV testing, prevention, and treatment resources, and other related resources, likely plays a key role in HIV risk for BMSM, as well as for MSM and other key populations in general. Despite these findings, there is a dearth of research on the role that social and structural-level factors, specifically neighborhood- or community-level HIV prevalence and access to resources close by in one’s neighborhood or community, play in explaining the incidence of HIV among BMSM.

In 2014, Raymond and his colleagues published a study in which they examined individual socioeconomic status, as well as community, social, and sexual network factors and their relationships to HIV risk among BMSM and White MSM (WMSM) living in San Francisco, CA. The results of the study indicated that BMSM were more likely than WMSM to live in higher HIV prevalence and lower income areas, as well as have higher numbers of serodiscordant partnerships (where the partners have differing HIV statuses), including potentially serodiscordant condomless sex acts. Additionally, while individual socioeconomic status was not associated with serodiscordant partnerships or serodiscordant condomless sex acts, area HIV prevalence was positively associated with serodiscordant condomless sex acts among BMSM. Raymond et al.’s findings form a foundation for the argument that future HIV prevention interventions must consider area HIV prevalence and HIV prevalence in social/sexual networks, and not just individual factors (e.g., individual behavioral change, poverty reduction).

While these studies offer important contributions to the field in terms of understanding whether and to what extent social and structural factors are associated with HIV risk for BMSM, there are a number of opportunities for researchers to build upon their work. For example, Raymond and his colleagues (2014) evaluated social and structural factors’ associations with HIV
risk behaviors (i.e., serodiscordant partnerships and serodiscordant condomless anal sex acts), rather than explicitly determining direct and indirect associations between social and structural factors and HIV acquisition. If there are statistically significant direct or indirect relationships between higher-level factors (e.g., HIV prevalence, lower average income in an area) and HIV seroconversion among BMSM, an even stronger argument for the importance of attending to higher level factors would be made.

**HIV Prevention: Relevant Social and Structural Factors**

Much of the extant literature concerning structural factors related to HIV outcomes focuses on the association between various social and structural variables and their association to HIV risk behaviors (e.g., serodiscordant condomless anal sex acts). However, equally important is whether there are associations between these social and structural variables and HIV prevention variables, such as, for example, individuals’ attitudes and behaviors concerning HIV prevention tools, like condoms, or, more recently, oral antiretroviral pre-exposure prophylaxis (PrEP).

**PrEP for HIV prevention and associated factors.** PrEP is a biomedical form of HIV prevention. PrEP is a once-daily oral antiretroviral medicine, composed of tenofovir disoproxil fumarate/emtricitabine (brand name Truvada®), taken by HIV negative persons who are at risk of HIV to prevent future infection. In clinical trials (e.g., Grant et al., 2010), PrEP has been shown to be a highly effective HIV prevention tool for communities at risk, as effective as condoms, when adherence is high; PrEP (with high adherence) reduces the risk of HIV transmission by over 90 percent (Grant et al., 2010). HIV scientists are hopeful about PrEP due to its ability to protect against HIV without penalizing sexual pleasure, and potentially even enhancing sexual pleasure (by reducing HIV-related anxiety; Calabrese & Underhill, 2015).
Despite the demonstrated efficacy of PrEP to protect against HIV transmission, HIV scientists know incredibly little about the actual implementation of PrEP, including awareness, access, and utilization, among communities that are most at risk for HIV (Matthews et al., 2016). For example, emerging evidence shows that relatively few BMSM who are eligible for PrEP are aware of it, and a small percentage report actually taking the medication (Eaton, Driffin, Bauermeister, Smith, & Conway-Washington, 2015), despite the fact that it has been approved by the U.S. Food and Drug Administration (FDA) since 2012. Some research has considered barriers to PrEP awareness, access, and use, and individual-level factors such as stigma (Calabrese & Underhill, 2015), low educational attainment, and lack of health insurance (Bauermeister, Meanley, Pingel, Soler, & Harper, 2013).

Much like the HIV risk literature, the HIV prevention literature focuses too narrowly on individual-level factors and tends to neglect social- and structural-level issues. While they have not yet been explored, social and structural contexts may have the potential to present barriers to PrEP uptake and adherence, particularly among vulnerable, disenfranchised communities, including BMSM. A study by Oldenburg et al. (2015) compared U.S. states by their number of policies against sexual orientation discrimination, as well as density of same-sex couples – factors which may represent each state’s level of stigma toward LGBT+ persons – and found that states with fewer policies against discriminatory sexual orientation practices and fewer same-sex couples was associated with decreased awareness of PrEP and increased discomfort discussing sexual health with healthcare providers among MSM. To date, no studies have compared geographic areas (i.e., states, counties, communities) regarding sentiment about PrEP (positive, negative, or neutral) and PrEP conversation “traffic” (i.e., frequency of conversations about PrEP).
The Potential Importance of Sentiment and Conversation “Traffic”

While a number of studies have explored potential PrEP users’ and their health care providers’ awareness of, and attitudes and opinions toward PrEP (e.g., Smith, Toledo, Smith, Adams, & Rothenberg, 2012; Tellalian, Maznavi, Bredeek, & Hardy, 2013), to date, no study has considered the general public’s PrEP awareness, attitudes and opinions. While understanding PrEP awareness and attitudes among potential users and potential prescribers is, no doubt, key, it is also important to understand the same from the general public, as their opinions may overall reflect or create a social climate that stigmatizes or otherwise hinders potential PrEP users’ willingness to use the medication to prevent HIV infection. A study by Farhat and colleagues (2016) of residents of high HIV prevalence neighborhoods in New York City found that only 25 percent of individuals were aware of PrEP, and over 50 percent of individuals endorsed negative stereotyped beliefs about PrEP users. Residents with low levels of homophobia, who were young, and who had LGBT+ friends were less likely to endorse negative beliefs and feelings about PrEP users. These findings indicate that there are quite varied attitudes and knowledge about PrEP among the general public, and that these attitudes and knowledge vary by demographic and interpersonal factors. Additional studies are warranted to determine whether PrEP attitudes – i.e., sentiment – and frequency of discussions (which, to some degree, indicates awareness) vary by geographic location (e.g., across cities) and are related to social- and structural-level factors.

Measuring General Public’s Sentiment and Discussions about PrEP.

Social media is an excellent resource for capturing the public’s thoughts, attitudes, and beliefs about a wide variety of topics, including PrEP. Harnessing the utility of social media, PrEP discussions and sentiment may be measured. Further, using social media from sites like Twitter where users often share their geographic location, social- and structural-level characteristics of
Twitter user’s locations (at, for example, the state or city level) can be examined with respect to their potential associations with the general public’s PrEP sentiment and frequency of PrEP conversations in each of those areas (e.g., states, cities).

**Approaching HIV Risk and HIV Prevention Research Questions with Social and Structural Variables**

In order to best determine whether a variety of social- and structural-level variables are associated with HIV/STI acquisition, as well as PrEP sentiment, the utilization of “big data” sources is necessary. Although there is no uniformly agreed upon definition, “big data” broadly refers to datasets that are “often characterized by their enormity and complexity” (Grant et al., 2012). These datasets are available in part because of affordable and easy-to-use technologies, including social media technologies, which have increased researchers’ abilities to generate large amounts of information (Young, 2015).

**Types of “Big Data”: Unstructured and Structured**

**Unstructured data.** Social media technologies offer a wealth of unstructured data (e.g., social media posts, messages, searches, and updates) that can be used to draw inferences about individuals’ real-time behaviors, attitudes, and sentiments about a wide variety of topics. Social media outlets that are frequently used for this purpose include Facebook and Twitter.

**Structured data.** Structured data includes public health surveillance data – or the “ongoing, systematic collection, analysis, and use of health-related data” (Sweeney et al., 2013, p. 558). Public health surveillance data have been used to better understand the epidemiology and transmission patterns of disease, as well as to prevent, and/or to control disease (including HIV). In the United States, state and local health departments are required by state law or regulation to report diagnoses of notifiable diseases, including HIV. Broadly, HIV surveillance,
according to Sweeney et al. (2013, p. 558), “describes the ‘who, what, when, and where’ descriptions of patterns of infection and disease occurrence that have guided public health prevention and control measures at both the population and individual levels.” In the past, HIV surveillance data have been used as a foundation for preventing new HIV infections and for monitoring and facilitating optimal HIV care for persons living with HIV (PLWH), including linkage, retention, and re-engagement in health care services to achieve the goal of virologic suppression (Bertolli et al., 2012; Evans, Van Gorder, Morin, Steward, Gaffney, & Charlebois, 2015). In addition, HIV surveillance data are used to estimate HIV morbidity and mortality and to detect populations or geographic areas in which HIV is highly concentrated (Hall et al., 2008). These estimates can serve as guides for decision-making around populations or communities that will benefit from increased efforts to scale up HIV prevention and treatment services (Evans et al., 2015). Websites like AIDSVu.org and HIVContinuum.org are particularly helpful in that they offer data at city, Zip Code, county, and state levels regarding how each area is doing to achieve HIV prevention and treatment goals outlined by the National HIV/AIDS Strategy (i.e., a five-year plan that details principles, priorities, and actions that guide the nation’s collective response to the epidemic) and by UNAIDS’ 90-90-90 treatment targets, which is that, by 2020, 90% of all people living with HIV will know their HIV status, 90% of people with a diagnosed HIV infection will receive sustained ART, and 90% of people receiving ART will be virally suppressed.

Unstructured and structured data both have the potential to offer important insights into factors relating to HIV/STI risk and acquisition, as well as public attitudes/sentiment toward emerging HIV prevention technologies, such as PrEP. Because the HIV epidemic among BMSM in the U.S. is so dire, paired with the fact that individual-level interventions simply are not enough
to control this epidemic alone, the use of these “big data” sources to inform where further studies, and allocation of additional resources, is most needed.

**Theoretical Orientation**

For the first two studies, a variety of social-ecological and social-epidemiological frameworks are used to guide the understanding of the roles of social and structural settings in HIV risk, HIV prevention, and HIV acquisition (Baral, Logie, Grosso, Wirtz, & Beyrer, 2013; Bronfenbrenner, 1992; Boerma & Weir, 2005; Latkin & Knowlton, 2005; Poundstone, Strathdee, & Celentano, 2004). For the present dissertation, social-ecological and social-epidemiological frameworks were used to guide the exploration of the relationships between HIV and STI infections among BMSM with community HIV prevalence, community socioeconomic deprivation, and access to HIV testing and prevention resources. These frameworks were also used for the third study, of which the primary purpose was to better understand the relationship between Twitter users’ sentiment toward PrEP for HIV prevention and various structural and social characteristics of their geographic areas.

In the first two studies, the health outcome of interest is HIV/STI infection. This outcome is at the core of Poundstone et al.’s (2004) social epidemiological framework and is Level 1 of said framework. The next layer, Level 2, represents the specific behaviors that put individuals at risk for HIV/STIs, such as condomless anal intercourse (CAI). Level 3 is composed of the individual-level factors (e.g., number of partners, income) that contribute to, or shape, HIV risk behaviors. The remaining two levels -- Levels 4 and 5 -- represent the larger, social and structural factors that influence HIV/STI risk behaviors and subsequent HIV/STI infections. The first two studies focus on Level 1 – HIV (or STI) infection – and factors at Level 4 (specifically, population density and neighborhood socioeconomic deprivation, both of which are social/Level 4 factors).
The third and final study focused on the relationship between Level 3, Level 4, and Level 5 factors by examining the content of conversations about pre-exposure prophylaxis (PrEP) for HIV prevention on social media (specifically Twitter) to determine users’ sentiment toward the HIV prevention tool. Twitter has a function where its users are able to “turn on” their location while they “tweet” (i.e., share information with their Twitter networks) or may share their location in their profiles. Therefore, for a fraction of the conversations about PrEP on Twitter, the authors’ actual or probable geographic location is known. Of primary interest in the third study is both what is said about PrEP and how often it is said in different geographic areas, and whether there are structural (Level 5) and social (Level 4) characteristics of those areas that are associated with the content and frequency of PrEP tweets originating from those areas.

**Use of Theory by Study**

The first study uses Poundstone et al.’s (2004) social epidemiological framework, as well as Boerma and Weir’s (2005) proximal-determinants framework, which is complementary to the social epidemiological framework, to understand the roles of social and structural factors in HIV acquisition. Like the social epidemiological framework (Poundstone et al., 2004), the proximal-determinants framework nests individual-level factors (including individual-level HIV risk factors) within social- and structural-level factors. Boerma and Weir’s (2005) model offers unique aspects beyond the social epidemiological framework in that it considers HIV risk behaviors (e.g., CAI) as proximal determinants of HIV infection, while other individual-level factors like demographic characteristics and psychosocial factors are considered underlying determinants. In their framework, social and structural factors are also considered to be underlying determinants that both directly and indirectly – predominately via HIV risk behaviors – influence HIV acquisition. Please refer to Figure 1 for a visual representation of Study 1’s theoretical framework.
The second study is guided by Poundstone et al.’s (2004) social epidemiological framework, and Earnshaw and Chaudoir’s (2009) HIV stigma framework. The HIV stigma framework explains various mechanisms of stigma and is complemented by the social epidemiological framework when questions around how individual-level and social- and structural-level factors interact with stigma to impact HIV acquisition and other HIV-related outcomes. We used Earnshaw and Chaudoir’s HIV stigma framework (2009) to guide our decision to examine the relationship between enacted stigma from health care providers and HIV/STI outcomes, and we employed Poundstone et al.’s (2004) social epidemiological framework to consider whether qualities of individuals’ neighborhoods directly and indirectly – by interacting with enacted stigma from health care providers – were associated with HIV/STI outcomes. Please refer to Figure 2 for a visual representation of Study 2’s theoretical framework.

The third study exclusively calls upon Poundstone et al.’s (2004) framework to explore whether Twitter users’ sentiment about PrEP for HIV prevention is related to social- and structural-level qualities of the areas in which they reside. Specifically, we examined whether sentiment and frequency of Twitter discussions about PrEP were related to HIV prevalence rate and proportion of the population with health insurance coverage in 40 different Metropolitan Statistical Areas in the U.S. Refer to Figure 3 for a visual representation of Study 3’s theoretical framework.

The Present Dissertation

This dissertation – divided into three separate but related studies, all of which use either structured or unstructured “big data” sources – is dedicated to (1) the study of area-level correlates of HIV/STI seroconversion among BMSM residing in the Atlanta, GA metropolitan and surrounding areas, and (2) exploring and sentiment toward, and frequency of discussing PrEP for
HIV prevention, and determining whether area-level social and structural factors are associated with sentiment and frequency of PrEP conversations.

The first study examines the link between Level 1 and Level 4 of Poundstone et al.’s (2004) framework. Using AIDSVu data, relationships between community HIV prevalence and access to HIV-related resources (including health care services) variables with testing HIV positive are explored. This study examines HIV seroconversion among BMSM specifically and considers community HIV prevalence and access to resources in the places that BMSM reside.

The second study uses Census data to create measures of community socioeconomic deprivation and population density to determine whether the level of socioeconomic deprivation in one’s community, or the population density of one’s community, are related to testing positive for HIV or other STIs.

The third and final study uses Twitter data to determine sentiment toward, and frequency of conversations about, PrEP. The sub-set of tweets for which a location is provided are mapped to visualize the geographic distribution of the tweets, and then metropolitan areas are compared based on their frequency of PrEP tweets and sentiment of PrEP tweets.
References


Figure 1

*Study 1’s Theoretical Framework*
Figure 2

*Study 2’s Theoretical Framework*
Figure 3

*Study 3’s Theoretical Framework*
Proximal and underlying determinants of HIV diagnosis among African-American/Black gay, bisexual, and other men and who have sex with men in the Southeastern United States
Chapter 2: Proximal and underlying determinants of HIV diagnosis among African-American/Black gay, bisexual, and other men who have sex with men in the Southeastern United States

The epidemiologic trends of human immunodeficiency virus (HIV) infection, paired with the adverse health outcomes that are associated with the disease, situate it as one of the most complex but critically important issues that has faced the public health field to date (Valdiserri, 2002). Since the 1980s, HIV has accounted for considerable morbidity and mortality both across the world and in the United States, with substantial financial and emotional costs to many individuals, families, and communities (Moyer, 2013). Over the past three decades, however, the U.S. HIV epidemic has evolved, such that it is now strongly concentrated in communities that are socially marginalized, stigmatized, and disenfranchised, including persons who are sexual orientation and/or racial or ethnic minorities (Pellowski, Kalichman, Matthews, & Adler, 2013).

African-American/Black persons in the U.S., including those who are gay, bisexual, and other men who have sex with men (i.e., BMSM) are heavily, and disproportionately, impacted by HIV (Centers for Disease Control and Prevention (CDC), 2017). A study of BMSM in six U.S. cities reported a HIV prevalence of 21 percent and a yearly HIV incidence rate of three percent (Koblin et al., 2013). Another study by Matthews and his colleagues (2016) found that, if current rates continue, an estimated 60 percent of BMSM will be HIV positive by the time they reach 40 years of age. Taken together, it is clear that the HIV epidemics among U.S. BMSM are serious and continuing issues that warrant immediate attention, resources, and energy for the development of innovative strategies to better understand and address these crises.

Social and structural contexts, which include the sociocultural and socioeconomic contexts of communities in which individuals reside, have contributed to the substantial burden of HIV
among BMSM (Levy et al., 2014). Based on the work of Latkin et al. (2010) and others, social and structural factors, which can be conceptualized as variables that exist outside the individual and are beyond the individual’s control, play distal but key roles in shaping health behaviors and outcomes. Boerma and Weir (2005) put forth the proximate-determinants framework, which organized determinants of HIV infection into one of two broad categories: proximal determinants and underlying determinants. Proximate determinants are behavioral and biological factors that directly place individuals at risk for acquiring HIV. Examples of proximate determinants include condomless anal and vaginal intercourse (CAI/CVI), concurrent sexual partners, injection drug use, and other sexually transmitted infections (STIs), among others. Underlying determinants, on the other hand, include individual-level demographic and psychosocial factors, including age, educational attainment, and employment, as well as HIV risk perceptions and perceived ability to communicate about HIV status. Qualities of social and structural contexts, including availability of resources and rate of new HIV diagnoses in a given area (e.g., a neighborhood), are also underlying determinants of HIV infection. In addition, Poundstone et al.’s (2004) social epidemiological framework was used, which aids in conceptualizing how factors at different levels are nested within one another and necessarily interact with one another. In Poundstone et al.’s (2004) framework, Level 1 represents the disease outcome (HIV and/or other STIs, in this case), which is nested in Level 2, which represents specific behaviors that put individuals at risk for HIV and/or other STIs, including drug- and sex-related risk behaviors (e.g., CAI). Level 2 is nested within Level 3, which represents individual-level factors (e.g., number of partners, income) that contribute to, or shape, HIV risk behaviors (e.g., CAI). Finally, Levels 4 and 5 of the social epidemiological framework represent social- and structural-level factors that influence HIV/STI outcomes. In the present study, variables at all five levels are examined. A visual representation
of the present study’s theoretical framework – which pulls from both Boerma and Weir as well as Poundstone – can be found in Figure 1.

The proximate-determinants framework and the social epidemiological framework are particularly useful for considering factors that may explain the elevated rates of HIV among BMSM, particularly those that extend beyond HIV risk behaviors. To date, most research concerning HIV risk among BMSM has considered either exclusively or primarily individual-level, proximal determinants (i.e., HIV sex- and drug-related risk behaviors) and less often have underlying determinants, particularly social and structural contexts, been considered. The lack of attention to factors outside of the individual is problematic, as there is evidence that, compared to their White counterparts, BMSM engage in comparable or lower rates of HIV risk behaviors (Millett et al., 2012). Eliminating racial disparities in HIV infection among BMSM will not be feasible without also addressing the structural barriers, which will necessitate the consideration of the social and structural composition of BMSM’s residential areas.

**Study Objectives**

The overarching focus of the current study was to gain an understanding of the roles of various qualities of social and structural settings in HIV acquisition among BMSM residing in the Atlanta, GA metropolitan and surrounding areas. The specific objectives of the study were to (1) examine the extent to which participants engaged in a variety of sexual and drug-related HIV risk behaviors; (2) assess the Zip Code HIV prevalence rate and availability of HIV testing, preventive care and substance abuse/treatment resources in the Zip Codes in which participants lived; (3) determine whether the Level 1, proximal (HIV risk behaviors) and/or underlying (Level 3 and 4 social and structural qualities) factors were related to HIV acquisition; (4) to investigate whether qualities of BMSM’s social and structural environments (i.e., Zip Code HIV prevalence rate,
availability of resources) interact with HIV risk behaviors with respect to HIV acquisition. With respect to the third and fourth aims, we developed the following hypotheses. We hypothesized that there would be direct, positive associations between Level 1 proximal determinants (HIV risk behaviors) and HIV acquisition; there would be direct, negative associations between accessibility of resources and HIV acquisition; there would be direct, positive associations between Zip Code HIV prevalence rate and HIV acquisition; and Zip Code HIV prevalence rate and accessibility of resources with HIV acquisition would each be moderated by HIV risk behaviors.

Methods

Sampling, Recruitment, and Enrollment

The present study, which was approved by University of Connecticut’s Institutional Review Board (IRB), was a part of a larger behavioral HIV prevention intervention trial, called Think Twice (Eaton et al., 2018). Think Twice participants (N=450) were BMSM for whom 12-month follow-up HIV test results are available.

Participants were recruited from gay and queer-friendly bars, clubs, bathhouses, parks, and street locations in the Atlanta, Georgia metropolitan and surrounding areas, as well as from online venues like classifieds (e.g., Craigslist), social media websites (e.g., Facebook), and websites and mobile applications for seeking sex partners and dates (e.g., Black Gay Chat, Jack’d).

In-person screening procedures included the use of electronic handheld devices (e.g., tablets), and over the phone screening utilized telephone screening software. For in-person screening procedures, recruiters approached individuals as they entered the above-mentioned target venues. Individuals were eligible to participate if they identified as Black/African-American; if they reported at least one act of CAI in the past year with a man; if they reported a HIV negative or unknown status; and if they were 18 years of age or older.
Individuals who met the eligibility criteria were scheduled for a baseline appointment at the research study site located in downtown Atlanta. At the baseline appointment, individuals were provided with written informed consent for all study procedures. Individuals were asked to take an HIV test (the OraQuick ADVANCE Rapid HIV-1/2 Antibody Test) to determine their HIV status, and those with unreactive tests were eligible for enrollment. Individuals with reactive tests were linked to care and were eligible for other studies.

Individuals who were enrolled in the study attended up to four in-person appointments (baseline, 3-month, 6-month, and 12-month follow-up appointments) at the study research site over a 12-month period. Each of the four study appointments included completing an assessment using Audio Computer Assisted Interviewing (ACASI) software. Finally, at the 12-month appointment, after their ACASI assessments were complete, participants were tested for HIV a second time.

**Measures**

Drawing upon the proximate-determinants framework proposed by Boerma and Weir (2005) and the social epidemiological framework from Poundstone et al. (2004), the factors associated with HIV seroconversion among BMSM, which are described below, were grouped into the following two broad categories: (1) *proximal determinants* and (2) *underlying determinants*. Proximal determinants included Level 2 HIV risk behaviors, while underlying determinants included Level 3 demographic characteristics and psychosocial factors, as well as Level 4 and 5 social and structural contextual factors at the Zip Code level.

**Dependent Variable**
The outcome of interest in the present study was HIV seroconversion. Participants whose OraQuick tests at the 12-month follow-up appointment were unreactive were coded as 0 (*HIV negative*), and participants whose tests were reactive were coded as 1 (*preliminary HIV positive*).

**Independent Variables: Proximal Determinants**

**HIV risk behaviors.** Items regarding HIV risk sex behaviors included the total number of receptive and penetrative CAI acts in the past three months, as well as the total number of CAI and condomless vaginal intercourse (CVI) acts in the past three months. In addition, participants reported HIV risk drug use behaviors, and were asked whether they had used the following drugs in the past three months: marijuana, crack, party drugs (i.e., cocaine, methamphetamine), or sex drugs (nitrite inhalants (i.e., ‘poppers’), sildenafil). Individuals who had not used these drugs in the past three months were coded as 0 (*no recent drug use*), while those who had used one or more of these drugs were coded as 1 (*recent drug use*). Participants were also asked about injection drug use (IDU), which was examined separately from the other drugs. Those who had not injected drugs in the past three months were coded as 0 (*no recent IDU*), and those who had injected drugs were coded as 1 (*recent IDU*).

**Independent Variables: Underlying Determinants (Individual-level)**

**Demographics.** Participants’ age, gender identity, sexual orientation, educational attainment, employment, and income were measured at the baseline appointment.

**HIV risk perceptions.** Participants were asked five questions (Kalichman, Eaton, Cain, Cherry, Pope, & Kalichman, 2006) at the baseline appointment regarding how much risk for HIV they perceived under certain sex behavior scenarios. Questions included “How risky is anal sex without a condom as the bottom partner with a man you just met who tells you his HIV status is negative?” Responses ranged from 0 = *no/low risk* to 10 = *very high risk*. These five
items were averaged to create one risk perception variable, and higher scores indicated greater perceived risk associated with CAI. This measure demonstrated good internal consistency (Cronbach’s $\alpha = .84$).

**HIV testing knowledge.** Participants were asked five items regarding their knowledge about HIV testing (Eaton et al., 2007). Questions included “I can be certain of my HIV test result even if I am having unprotected sex around the time of the test.” Response set included *no* or *yes* and correct answers were given a point and incorrect answers were given no points. The answers were summed and therefore higher scores indicated a greater level of HIV testing knowledge.

**Independent Variables: Underlying Determinants (Social- and Structural-levels)**

**Zip Code-level HIV prevalence rate.** The HIV prevalence rate (per 100,000) for the Zip Codes in which participants lived was examined because of its importance in estimating the probability of potential exposure to HIV. These data were acquired from AIDSVu.org and reflect HIV prevalence among the adult/adolescent population within each county/ZIP-code in 2015.

**Distance to HIV testing and health care services.** Service factors included distance to (1) the nearest HIV testing service site, (2) the nearest PrEP provider site, and (3) the nearest health center. These data were gathered from the HIV.gov Services Locator website at [https://locator.aids.gov/](https://locator.aids.gov/). Distance (in miles) was calculated from each participants’ home address to the nearest service of each of the four above-mentioned services.

**Dependent Variables**

**HIV and STI diagnoses.** The first variable, HIV status, was captured at the 12-month follow-up appointment, participants were tested for HIV. Response set included 0 = *unreactive* (HIV negative) or 1 = *reactive* (preliminary HIV positive).

**Statistical Analyses**
HIV risk behaviors, individual-level demographic and psychosocial characteristics, and area-level qualities of social and structural contexts were summarized as Ns and percentages or with Ms and SDs. Odds ratios (ORs) were employed to test whether individual-level and area-level characteristics differed significantly between HIV positive and HIV negative participants.

To account for the “nesting” of BMSM within Zip Codes, we fit a total of four random intercept logistic regression models to examine the relationships between proximate (i.e., HIV risk behaviors) and underlying (i.e., demographic and psychosocial, as well as social and structural contextual) determinants with HIV seroconversion among BMSM. We assessed the direct and moderating effects of these factors by fitting four sequential models, each examining a set of variables (and, in the fourth model, interactions between variables) that conceptually belong to each level of the proximate-determinants model (i.e., proximal determinants (HIV risk behaviors) and individual-level and social/structural-level underlying determinants). In Model 1, the baseline model, proximate determinants, including all HIV sex- and drug-related risk behaviors, were included (i.e., number of receptive and insertive condomless anal intercourse acts, number of condomless vaginal intercourse acts, IDU, and other drug use). In Model 2, individual-level demographic and psychosocial underlying determinants were added (i.e., age, sexual orientation, educational attainment, employment, income, HIV risk perceptions, and HIV testing knowledge). In Model 3, social and structural contextual underlying determinants were included (i.e., HIV prevalence in Zip Code, availability of HIV testing services, health care centers, and PrEP providers). Model 4 included interactions between any statistically significant social and structural determinants and any statistically significant HIV risk behaviors.

All analyses were performed using RStudio with the lme4 and simr libraries. Less than 5% of data were missing for any given variable included in the above-mentioned analyses, and an α
level less than .05 was considered statistically significant, and an $\alpha$ level of .10 to .05 was considered to be a trending relationship.

**Statistical Power and Effect Size Considerations**

To estimate effect sizes for the community-level variables in our fixed-effects models, R Studio’s lme4 and simr libraries were employed to perform automated power analysis simulations (Gelman & Hill, 2007; Bolker, 2008). The simr package for R in particular provides tools that make setting up and running power analysis simulations (in this case, Monte Carlo simulations) for fixed-effects and mixed-effects models relatively straightforward, particularly when said models are performed in lme4.

A power analysis simulation requires multiple steps, which are in this case automated by the simr package. The steps include (1) generation of multiple simulated datasets, (2) refitting the model to subsets of the new data, (3) applying statistical tests to the fitted models, and finally, (4) collating and reporting on the results. The simr library performs these steps automatically and can calculate an effect size for your variable(s) in question with a few lines of R code.

The simr package allows for specified modification of parameters as well, such that power to detect a variety of trends can be calculated. It is worth noting, however, that retrospective ‘observed power’ calculations, where the target effect size comes from the data, do typically give misleading results (Green & MacLeod, 2016); as such, the following results should be interpreted with caution.

Typically, with observed power calculations, scientists can either choose an effect size slightly smaller than the one that was observed, or simply test their observed effect size against the null model by modifying the variable(s) in question’s effect size to 0. The second option was chosen for the present study, as typically choosing an effect size requires some prior knowledge
of what the specified effect size would be; given that these variables have never before been examined, we chose to test our observed effect sizes against the null model.

The simr package provides output with power for the variable in question, based on 1,000 simulations and $N$ equivalent to the actual dataset (in this case, 549). Power analysis simulations were performed for the four area-level variables, which constitute the heart of the primary purpose for the present study (Zip Code HIV prevalence rate, and distance to the nearest HIV testing site, PrEP provider, and health care center. Power analysis results were reported in percentages and reflected the likelihood of rejecting the null hypothesis when it is false. Simulated power (likelihood of rejecting the null hypothesis when it is indeed false) for Zip Code HIV prevalence rate was 27.6% (95 CI: 0% to 38%), for distance to nearest HIV testing site was 23.4% (95 CI: 0% to 38%), for distance to nearest PrEP provider was 9.0% (95 CI: 0% to 38%), and for distance to nearest health care center was 29.1% (95 CI: 0% to 38%). These percentages indicate that the study was under-powered to detect an effect; however, using a combination of HIV and STI diagnoses as the outcome variable did not substantially increase power, and effects were still observed with HIV status as the outcome variable.

**Results**

**HIV Risk Behaviors**

Overall, the sample engaged in approximately 2.47 acts of receptive CAI in the past three months, and 3.65 acts of insertive CAI in the past three months. During that same time frame, the overall sample engaged in an average of 1.13 acts of CVI. Participants whose HIV tests at the 12-month follow-up appointment were preliminary positive (hereto referred to simply as HIV positive) did not significantly differ from those who tested HIV negative at the 12-month follow-
up appointment on any of the above-mentioned variables. Additional information about each of the above-mentioned HIV risk behaviors can be found in Table 1.

**Demographics**

The overall sample \((N=450)\) was, on average, was somewhat younger in age \((M = 33.61, SD = 11.78, \text{range} = 18-62)\). Most participants identified as gay or same gender loving (45.30%), followed by bisexual (39.60%) and heterosexual (15.10%). It is worth reiterating that, while some individuals identified as heterosexual, to be eligible for this study, all individuals must have engaged in CAI (either insertive or receptive CAI) at least one time in the past three months with a man. Participants who tested HIV positive were significantly younger in age \((OR = 0.90, 95\% \text{CI: 0.85-0.96, } p = 0.002)\) than those who tested HIV negative, however, there were no statistically significant differences between participants who tested HIV positive and those who tested HIV negative with respect to gender identity or sexual orientation.

While the majority of the sample reported receiving a college education (61.1%), most of the sample (44.7%) reported being unemployed. The remainder of the sample reported working (36.2%), being on disability (6.0%), being a student (8.90%), or other (4.20%). The majority of the sample had annual incomes equal to or less than $20,000 (74.71%). There were no statistically significant differences between individuals who tested HIV positive and those who tested HIV negative with respect to education, employment, or income. Information about each of the above-mentioned demographic variables can be found in Table 2.

**HIV Risk Perceptions**

The overall sample’s HIV risk perceptions were relatively high \((M = 7.33, SD = 1.72, \text{range} = 0-9)\). Participants who tested HIV positive were trending toward having lower HIV risk
perceptions, but the association was not statistically significant (OR = 0.97, 95% CI: 0.84-1.01, p = 0.118). Information about HIV risk perceptions can be found in Table 2.

**HIV Testing Knowledge**

Overall, the sample had average knowledge of HIV testing ($M = 3.41$, $SD = 1.22$, range = 0-5). There were no statistically significant differences between those who tested HIV positive and those who tested HIV negative with respect to HIV testing knowledge. Additional regarding HIV testing knowledge can be found in Table 2.

**Zip Code HIV Prevalence Rate**

The average HIV prevalence rate in the Zip Codes represented by the present sample was $M = 1,009.71$ (per 100,000; $SD = 749.44$, range = 30-2689). This HIV prevalence rate is 1.79 times that of the HIV prevalence rate for the state of Georgia. Participants who tested HIV positive at the 12-month follow-up did not differ from those who tested HIV negative with respect to Zip Code-level HIV prevalence. Information regarding Zip Code HIV prevalence rate can be found in Table 2.

**Distance to Nearest HIV Testing Site, PrEP Provider, and Health Care Center**

The average distance participants’ residences were to the nearest HIV testing site, in miles, was 1.61 ($SD = 1.65$, range = 0.10-12.72). The average distance from participants’ residences to the nearest PrEP provider, on average, was 5.66 miles ($SD = 5.92$, range = 0.10-37.70), and the average distance to the nearest health care center was 2.33 miles ($SD = 2.56$, range = 0.10-17.0). Participants who tested HIV positive at the 12-month follow-up appointment were significantly further from the nearest HIV testing site (OR = 1.21, 95% CI: 1.01-1.47, $p = 0.05$). Individuals who tested HIV positive were trending toward being further away from the nearest PrEP provider (OR = 1.04, 95% CI: 0.98-1.11, $p = 0.11$) and being further away from the nearest health care center.
(OR = 1.11, 95% CI: 0.99-1.25, p = 0.08), though these relationships did not reach statistical significance. Information regarding distance to nearest HIV testing site, PrEP provider, and health care center can be found in Table 2.

**HIV Diagnoses**

A total of n=25 participants seroconverted during the study (i.e., tested positive for HIV at the 12-month follow-up appointment). This rate of seroconversion (4.55%) is alarmingly high but similar to that of other BMSM cohorts residing in the U.S. (Matthews et al., 2016). The remaining n=425 participants tested HIV negative at the 12-month follow-up appointment.

**Multilevel Random Intercept Logistic Regression Models**

For all models, Model 1 through Model 4, the dependent variable was HIV status. Because HIV status is a binary variable, we employed a binomial distribution and the logit function for our analyses, and we included in each of the four models a random intercept. The grouping variable used was Zip Code in which each participant resided, of which there were N=84 represented.

In Model 1, which included the proximal determinants (i.e., HIV risk behaviors), including both receptive and insertive CAI, as well as CVI and drug use, receptive CAI was the only variable that was significantly associated with HIV status. The relationship between receptive CAI and HIV status was positive, such that an increase in the number of receptive CAI acts in the past three months increased the probability of testing positive for HIV at the 12-month follow-up ($\beta = 0.879, SE = 0.347, p = 0.012$). Model 1 results can be found in Table 3.

In Model 2, which included any statistically significant variables from Model 1 (in this case, only receptive CAI), as well as individual-level underlying determinants (i.e., age, sexual orientation, educational attainment, employment status, income, HIV risk perceptions, and HIV testing knowledge). The only variables that were statistically significantly related to the outcome
variable, HIV status, were age and HIV risk perceptions. Specifically, there was a significant, negative relationship between age and HIV status, such that as age decreased, the likelihood of testing HIV positive at the 12-month follow-up increased ($\beta = -0.002, SE = 0.001, p = 0.029$). In addition, HIV risk perceptions were significantly, positively related to testing HIV positive ($\beta = 0.387, SE = 0.176, p = 0.028$). All other variables were not statistically significantly related to the outcome variable. Model 2 results can be found in Table 3.

In Model 3, age remained significantly, negatively associated with HIV status ($\beta = -0.005, SE = 0.001, p = 0.004$) and HIV risk perceptions remained significantly, positively associated with HIV status ($\beta = 0.554, SE = 0.474, p = 0.016$). In addition, three of the four area-level variables were significantly, positively associated with testing positive for HIV. Specifically, as participants’ Zip Code HIV prevalence rate increased, likelihood of testing positive for HIV increased ($\beta = 1.341, SE = 0.513, p = 0.010$). Finally, as distance to nearest HIV testing site ($\beta = 0.964, SE = 0.348, p = 0.007$) and health care center ($\beta = 1.061, SE = 0.318, p = 0.002$) increased, likelihood of testing HIV positive increased. All other variables were not statistically significantly related to the outcome variable. Model 3 results may be found in Table 3.

Finally, in Model 4, HIV risk perceptions remained statistically significant and positively related to HIV status ($\beta = 0.002, SE = 0.001, p = 0.021$). Zip Code-level HIV prevalence rate was significantly, positively associated with testing HIV positive ($\beta = 0.962, SE = 0.356, p = 0.007$), as was distance to nearest HIV testing site ($\beta = 1.093, SE = 0.332, p = 0.002$) and distance to nearest health care center ($\beta = 1.993, SE = 0.271, p < 0.001$).

Additionally, in Model 4, results revealed that there were three two-way interactions that were significant. The interactions between receptive CAI, the only (and epidemiologically primary) significant risk factor for HIV seroconversion, and the three statistically significant area-
level variables (i.e., Zip Code HIV prevalence rate, distance to nearest HIV testing site, and distance to nearest health care center) were included. First, receptive CAI × Zip Code HIV prevalence rate was significant and positive, such that as Zip Code HIV prevalence increased, the relationship between number of receptive CAI acts and testing HIV positive at the 12-month follow-up appointment strengthened ($\beta = 0.971, SE = 0.294, p = 0.001$). Second, receptive CAI × distance to nearest HIV testing site was significant and positive, such that as distance to nearest HIV testing site increased, number of receptive CAI acts and testing HIV positive became stronger ($\beta = 0.909, SE = 0.332, p = 0.008$). Finally, receptive CAI × distance to nearest health care center was significant and positive, such that as distance to nearest health care center increased, number of receptive CAI acts and testing HIV positive became stronger ($\beta = 0.996, SE = 0.209, p < 0.001$).

Results from Model 4 are shown in Table 3.

Model fit was measured using the following deviance statistics: -2 log likelihood (-2LL), AIC, and BIC, which all indicated that model fit improved from Model 1 to Model 4, and that model fit was best in Model 4. In Model 4, -2LL was 27.6, AIC was -25.6, and BIC was -22.2. Deviance statistics for all models are reported in Table 3.

**Discussion**

The present study’s findings, arguably, are both in line with, and substantially expand upon the findings of, previous studies in this area. Namely, we found that, the most robust proximal, behavioral risk factor for HIV infection among BMSM was, in fact, frequency of receptive CAI acts, which supports prior indications that this particular behavioral practice carries the highest per act risk of all HIV-related risk behaviors (Boily et al., 2009; Patel, Borkowf, Brooks, Lasry, Lansky, & Mermin, 2014). In Models 2 through 4, however the main effect of frequency of receptive CAI acts was not statistically significant, suggesting that, after controlling for underlying
determinants at the individual- and area-levels, frequency of receptive CAI is no longer associated with HIV seroconversion among BMSM. Our results echo the cries of prior findings in the HIV disparities literature – such as the fact that BMSM are five times more likely to be HIV infected than their White counterparts despite similar levels of engagement in CAI (Millett et al., 2012) – which request allocating more attention to underlying determinants of infection.

While frequency of receptive CAI acts had no statistically significant main effect in Models 2 through 4 after the addition of the underlying determinants of infection, the variable did, however, significantly interact with three of the four proposed area-level variables, including distance to nearest HIV testing site, distance to nearest health care center, and Zip Code HIV prevalence rate. This finding is in line with the proximate-determinants framework (Boerma & Weir, 2005), which argues that area-level determinants affect HIV infection outcomes through proximate, HIV risk behavior determinants. This was certainly true in the present study, and significant main effects were observed for these area-level variables as well. These findings also provide additional support for the notion that area-level variables must not be ignored by HIV prevention scientists if we are to develop and administer the most targeted, informed, and effective HIV prevention interventions for BMSM and, likely, for other key populations.

Prior studies that have ventured beyond proximate, HIV risk behavior determinants of HIV infection among BMSM have largely considered individual-level underlying determinants, including demographic characteristics and psychosocial characteristics, such as age, income, as well as risk reduction intentions, condom use norms, and AIDS conspiracy beliefs (Kelly et al., 2016), as well as clinically significant mental health concerns (Reisner et al., 2009), and being “on the down low” (Millett et al., 2005). However, no study to date has considered whether the way in which BMSM’s residential environments are composed, both socially and structurally, facilitate
risk. As such, exploring these variables’ relation to HIV seroconversion among BMSM is both novel and acts as a step toward completing the larger social-ecological “picture” and deepening our understanding of HIV risk for members of this community.

The present study’s primary finding was that frequency of receptive CAI acts was more strongly associated with testing HIV positive for BMSM who lived in Zip Codes with higher HIV prevalence rates, as well as for BMSM who lived further away (in miles) from their nearest HIV testing site and from their nearest health care center. These findings may indicate that BMSM’s sexual networks are relatively geographically close to their place of residence. If this is indeed the case, then it is plausible that HIV prevalence rate in one’s Zip Code to some degree reflects the density of HIV in one’s sexual network, and therefore one’s likelihood of exposure to HIV. The findings indicate that, in areas where there there may be a high likelihood of exposure to HIV, risk behavior becomes more strongly associated with infection. While no study has before tested this relationship, a study of MSM in San Francisco (Das et al., 2010) did find that Zip Code-level viral load and HIV prevalence did predict new infections, which supports the present study’s findings.

In a time where geospatial dating and hook-up applications are nearly ubiquitously used to seek sex partners (and, indeed, they recommend partners in close geographic proximity; Eaton, Maksut, Gamarel, Siembida, Driffin, & Baldwin, 2016), it is quite possible that some (or even many) BMSM have at least somewhat geographically close sexual networks. In addition, prior studies indicate that BMSM are significantly more likely than MSM of other races/ethnicities to have same-race partnerships (Berry, Raymond, & McFarland, 2007; Millett et al., 2012). The finding that greater distance to nearest preventive and testing services was associated with higher likelihood of HIV infection seems plausible for two reasons. First, that distance (i.e., difficulty of accessibility) is the same for other MSM and BMSM in their geographic area, which may indicate
a higher number of unrecognized infections in that area, and in their networks. Second, it may indicate lower level of knowledge about HIV and HIV prevention, if distance to services that are a major provider of said information, is far and therefore less accessible. Lower knowledge of HIV and how to prevent HIV may mediate the moderated relationship between distance to services and frequency of receptive CAI acts being associated with testing positive for HIV.

Both higher HIV prevalence rates and greater distance to needed services may indicate poor access to preventive and regular care overall in an area, both for the BMSM in the present study themselves and also potentially for the members of their sexual networks (who may or may not also be BMSM). This could lead to higher unrecognized (and therefore untreated) HIV infection in one’s community, and also potentially within one’s sexual network, therefore increasing one’s likelihood of exposure to HIV, and to acquiring the disease. There is evidence that undiagnosed infections are more common amongst BMSM than their White counterparts (Millett et al., 2012), and this study may help to understand the larger picture as to why that may be the case, and part of how to address it (i.e., increasing geographic accessibility of services). Eliminating disparities in HIV infection and other related outcomes requires better access to care (Tobias, Cunningham, Cunningham, & Pounds, 2007), of which geographic accessibility is an important part.

**Redefining HIV Risk**

The present study’s findings push us to question whether our current conceptualizations and quantifications of “HIV risk” are acceptable, particularly for marginalized communities such as BMSM. What does it mean to be “at high risk” for HIV infection, and how do – and should – scientists systematically identify and service individuals who have the highest probability of future HIV infection? The findings of the present study urge us to abandon the practice of conceptualizing
particular people as “high risk” individuals based on type and frequency of their sex practices, and also urge us to move away from attempting to address only those practices in prevention interventions. Instead, HIV prevention scientists must consider how these practices operate within, and differ between, particular contexts. In other words, receptive CAI is not always a “high risk” behavior; the results of the present study suggest it depends on circumstance. Considering receptive CAI to be a “high risk” behavior regardless of circumstance essentially suggests that individuals who engage in receptive CAI are, across the board, “high risk” individuals, and HIV risk can be minimized or eliminated if only individuals choose to discontinue “risky behaviors” such as receptive CAI. This is not an empathetic, sex positive, or acceptable approach to addressing HIV epidemics among BMSM, or among any population.  

The reconceptualization of “HIV risk” as something for which individuals are not fully personally responsible is critical for a number of reasons. First, studies show that attributions of responsibility lead to increased anger, decreased empathy, and diminished willingness to help others in need; as such, attributions of personal responsibility for HIV infection has the potential to have profound influence on service and policy responses to the epidemic (McDonell, 1993). In addition, without addressing factors external to the individual (i.e., the root causes or structures that affect individual risk or vulnerability to HIV), HIV epidemics will continue to persist (Gupta, Parkhurst, Ogden, Aggleton, & Mahal, 2008).  

Instead, more holistic, ecologically based, balanced strategies to address HIV epidemics, and to promote health, among key affected populations such as BMSM are needed. The present study revealed that area-level variables, including distance to one’s nearest HIV testing site, distance to one’s nearest PrEP provider, and distance to one’s nearest health center, and HIV prevalence in the area in which one resides, play important roles in HIV acquisition among BMSM.
BMSM who lived in neighborhoods with higher HIV prevalence, and who were further away from
the above-mentioned services, were more likely to test positive for HIV after controlling for known
HIV risk behavior and individual-level determinants of infection.

Limitations

The findings of the present study must be interpreted in light of its limitations. First and
foremost, our sample had $n=25$ BMSM test positive for HIV at the end of the study. While this is
an epidemiologically significant number, statistically, it left our analyses under-powered. Still, we
were able to detect effects in three of the four area-level variables; despite this, we may have
observed an effect for distance to PrEP provider had we more power. In the future, such analyses
should be performed with larger datasets with similar or higher seroconversion rates. Secondly,
the study consisted of variables from multiple sources and from reasonably similar, but still
different, time points. Data from the HIV risk behavior- and individual-levels were self-report, and
were therefore subject to social desirability bias. Participants were recruited from gay-friendly
establishments and from gay-friendly websites and dating apps, which may, to some degree, skew
the representation of BMSM in this study. Data for area-level variables were at the Zip Code level,
and a more nuanced analysis with a smaller unit of analysis would have been preferable as Zip
Codes span a large area in some cases, had AIDSVu offered data at that level.

Conclusions

The present study offered important insights into the relevance of area-level variables and
the role that they play in BMSM’s HIV seroconversion. We found that area-level risks (i.e., higher
HIV prevalence rate, and further distance from needed services) were associated with testing
positive for HIV among our sample. We also found that adding underlying determinants of HIV
infection bumped any HIV risk behaviors (i.e., proximal determinants) out of statistical
significance. Taken together, these findings suggest that these area-level variables play an important role in HIV infection for BMSM. As such, interventions must include and address issues at these levels (Gupta et al., 2008) by better allocating resources to areas in greatest need, and improving basic level accessibility of services via transportation offerings and additional services being placed in low-resourced areas.
References


systems model for considering structural factors in HIV prevention and detection. *AIDS and Behavior, 14*(2), 222-238.


connection: The importance of engagement and retention in HIV medical care. *AIDS Patient Care and STDs*, 21(S1), S-3.

Table 1

*HIV Risk Behaviors by HIV Status at the 12-month Follow-up Appointment*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>HIV Negative (n = 425)</th>
<th>HIV Positive (n = 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Number of receptive CAI acts in the past 3 mo.</td>
<td></td>
<td>2.29 (7.03)</td>
<td>2.91 (3.20)</td>
</tr>
<tr>
<td>Number of insertive CAI acts in the past 3 mo.</td>
<td></td>
<td>3.68 (8.11)</td>
<td>0.96 (1.52)</td>
</tr>
<tr>
<td>Number of CVI acts in the past 3 mo.</td>
<td></td>
<td>1.10 (5.25)</td>
<td>0.08 (0.41)</td>
</tr>
<tr>
<td>Injection drug use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td>8 (1.9)</td>
<td>2 (8.0)</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>497 (98.1)</td>
<td>23 (92.0)</td>
</tr>
<tr>
<td>Other drug use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td>323 (63.8)</td>
<td>20 (80.0)</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>182 (36.2)</td>
<td>5 (20.0)</td>
</tr>
</tbody>
</table>
### Table 2

*Individual-level and Area-level Factors by HIV Status at the 12-month Follow-up*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>HIV Negative</th>
<th>HIV Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(n = 425)</td>
<td>(n = 25)</td>
</tr>
<tr>
<td></td>
<td>(N (%))</td>
<td>(N (%))</td>
<td></td>
</tr>
<tr>
<td>Sexual orientation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same gender loving, gay, or homosexual</td>
<td>188 (44.4)</td>
<td>14 (56.0)</td>
<td></td>
</tr>
<tr>
<td>Bisexual</td>
<td>172 (40.7)</td>
<td>9 (36.0)</td>
<td></td>
</tr>
<tr>
<td>Straight or heterosexual</td>
<td>63 (14.9)</td>
<td>2 (8.0)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>31 (7.3)</td>
<td>1 (4.0)</td>
<td></td>
</tr>
<tr>
<td>High school or GED</td>
<td>124 (29.2)</td>
<td>10 (40.0)</td>
<td></td>
</tr>
<tr>
<td>College or higher</td>
<td>270 (63.5)</td>
<td>14 (56.0)</td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>182 (42.8)</td>
<td>10 (40.0)</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>158 (37.2)</td>
<td>13 (52.0)</td>
<td></td>
</tr>
<tr>
<td>On disability</td>
<td>29 (6.8)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Student</td>
<td>42 (9.9)</td>
<td>1 (4.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>14 (3.3)</td>
<td>1 (4.0)</td>
<td></td>
</tr>
<tr>
<td>Annual income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0-$19,999</td>
<td>311 (74.0)</td>
<td>19 (76.0)</td>
<td></td>
</tr>
<tr>
<td>$20,000 and above</td>
<td>109 (26.0)</td>
<td>6 (24.0)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>34.14 (11.86)</td>
<td>25.38 (6.33)</td>
<td></td>
</tr>
<tr>
<td>HIV risk perceptions</td>
<td>7.33 (1.72)</td>
<td>7.63 (1.31)</td>
<td></td>
</tr>
<tr>
<td>HIV testing knowledge</td>
<td>3.70 (1.26)</td>
<td>3.87 (1.03)</td>
<td></td>
</tr>
<tr>
<td>Zip Code HIV prevalence rate</td>
<td>868.18 (619.69)</td>
<td>1009.44 (756.30)</td>
<td></td>
</tr>
<tr>
<td>Distance to nearest HIV testing site</td>
<td>1.58 (1.67)</td>
<td>2.33 (2.35)</td>
<td></td>
</tr>
<tr>
<td>Distance to nearest PrEP provider</td>
<td>5.62 (5.88)</td>
<td>7.39 (8.17)</td>
<td></td>
</tr>
<tr>
<td>Distance to nearest health care center</td>
<td>2.43 (2.63)</td>
<td>3.48 (3.65)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3

**Fixed-effects Estimates for Models 1-4 with HIV Status as the Outcome Variable**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β (SE)</td>
<td>p value</td>
<td>β (SE)</td>
<td>p value</td>
</tr>
<tr>
<td>Intercept</td>
<td>-.012 (.309)</td>
<td>.968</td>
<td>.098 (.266)</td>
<td>.713</td>
</tr>
<tr>
<td>Proximate Determinants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of receptive CAI acts in the past 3 mo.</td>
<td>.879 (.347)</td>
<td>.012*</td>
<td>.421 (.113)</td>
<td>.651</td>
</tr>
<tr>
<td></td>
<td>.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of insertive CAI acts in the past 3 mo.</td>
<td>.099 (.162)</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>.542</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of CVI acts in the past 3 mo.</td>
<td>.033 (.217)</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>.878</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injection drug use</td>
<td>-.020 (.028)</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>.471</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underlying Determinants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual orientation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same gender loving, gay, or homosexual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bisexual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Straight or heterosexual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or GED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College or higher</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On disability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0-$19,999</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$20,000 and above</td>
<td>-.002 (.001)</td>
<td>.029*</td>
<td>-.005 (.001)</td>
<td>.004**</td>
</tr>
<tr>
<td></td>
<td>.086</td>
<td></td>
<td>.001 (.001)</td>
<td>.086</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-.002 (.001)</td>
<td>.029*</td>
<td>-.005 (.001)</td>
<td>.004**</td>
</tr>
<tr>
<td>HIV risk perceptions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.387 (.176)</td>
<td>.028*</td>
<td>.554 (.474)</td>
<td>.016*</td>
</tr>
<tr>
<td></td>
<td>.385</td>
<td></td>
<td>.002 (.001)</td>
<td>.021*</td>
</tr>
<tr>
<td>HIV testing knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-.034 (.039)</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zip Code HIV prevalence rate</td>
<td>-</td>
<td>-</td>
<td>1.341 (.513)</td>
<td>.962 (.356)</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---</td>
<td>---</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Distance to nearest HIV testing site</td>
<td>-</td>
<td>-</td>
<td>.964 (.348)</td>
<td>1.09 (.332)</td>
</tr>
<tr>
<td>Distance to nearest PrEP provider</td>
<td>-</td>
<td>-</td>
<td>.217 (.410)</td>
<td>-</td>
</tr>
<tr>
<td>Distance to nearest health care center</td>
<td>-</td>
<td>-</td>
<td>1.061 (.318)</td>
<td>1.99 (.271)</td>
</tr>
</tbody>
</table>

**Interactions**

<table>
<thead>
<tr>
<th>Receptive CAI × Zip Code HIV prevalence rate</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>.971 (.294)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptive CAI × Distance to nearest HIV testing site</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.909 (.332)</td>
</tr>
<tr>
<td>Receptive CAI × Distance to nearest health care center</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.996 (.209)</td>
</tr>
</tbody>
</table>

**Deviance Statistics**

<table>
<thead>
<tr>
<th>-2 log likelihood</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>79.8</td>
<td>81.8</td>
<td>51.6</td>
</tr>
<tr>
<td>48.1</td>
<td>50.0</td>
<td>51.6</td>
</tr>
<tr>
<td>-6.67</td>
<td>-4.66</td>
<td>-.732</td>
</tr>
<tr>
<td>-27.6</td>
<td>-25.6</td>
<td>-22.2</td>
</tr>
</tbody>
</table>

Note. * * p < .05, ** p < .01, p < .001. *Smaller is better when comparing models.
Neighborhood poverty, enacted stigma from health care providers, and HIV/STI among Black gay, bisexual, and other men who have sex with men in Atlanta, Georgia
Chapter 3: Neighborhood factors, stigma, and HIV/STIs among Black men who have sex
with men in Atlanta, Georgia

In 2014, Vaughan, Rosenberg, Shouse, and Sullivan published a seminal study which illustrated the importance of evaluating social and structural factors in understanding health outcomes, specifically, human immunodeficiency virus (HIV) acquisition. In their study, the authors showed that the well-known racial/ethnic disparities in HIV acquisition between African-American/Black gay, bisexual, and other men who have sex with men (BMSM) and their White counterparts (such that BMSM, despite similar levels of engagement in HIV risk behaviors, are approximately five times more likely to acquire HIV; Hall et al., 2008) was explained by poverty only in highly populated, urban centers. In rural areas, the HIV infection disparity between BMSM and White MSM (WMSM) remained significant, even after controlling for poverty. Vaughan and his colleagues concluded that the association between incident HIV infections and poverty varied by levels of urbanization; however, they failed to offer an alternative explanation for potential mechanisms of HIV infection among BMSM in more rural areas.

There is evidence that stigma may be one of the potential factors that places BMSM residing in rural areas at elevated risk for HIV, as well as other sexually transmitted infections (STIs). According to Kalichman, Katner, Banas, and Kalichman (2017), stigma is a widely experienced phenomenon; however, stigma is not geographically uniformly distributed. In their study, Kalichman and his co-authors found that, in rural settings, as compared to more urban areas, stigma may be more prevalent and more intensely experienced. Other studies have found that stigma may prevent persons from adequately engaging in healthcare by interfering with transportation, raising concerns about confidentiality, and inhibiting one’s ability to feel free to talk openly with his or her healthcare provider (Pellowski et al., 2013). Taken together, there is
evidence that individuals residing in rural areas who experience stigma – including, potentially, BMSM – may be at particularly high risk for HIV/STI acquisition and transmission.

Neither Vaughan et al. (2014) nor Kalichman et al. (2017) considered the role that stigma on behalf of racial/ethnic and sexual orientation minority identities may play in acquisition of HIV and other STIs. While Kalichman et al. (2017) considered how and to what degree AIDS stigma might impact HIV positive individuals’ health outcomes in rural settings, the links between racial/ethnic and sexual orientation minority stigma, population density (or rurality), and HIV/STI acquisition remains unevaluated. One study conducted with rural MSM found that general stigma emanating from families, healthcare providers, and rural communities was strongly associated with mental health outcomes and HIV risk behavior (Preston, D’Augelli, Kassab, & Starks, 2007), both of which are factors that have demonstrated links, both proximal and distal, to HIV/STI acquisition.

Understanding the link between racial/ethnic and sexual orientation minority stigmas and population density as it relates not only to HIV, but also to other STIs, is critical for BMSM in particular. STIs are among the few individual-level factors that have been shown to partially explain BMSM’s elevated rates of HIV infection as compared to MSM of other races/ethnicities (Millett et al., 2012). Furthermore, the biological link between STIs and subsequent HIV infection has been well documented (Ward & Rönn, 2010; i.e., STIs cause direct mucosal disruption, the presence of HIV target cells in the genital treatment, and increased HIV load in plasma and genital secretions). In addition, previous studies have confirmed a relationship between STI diagnoses and higher rates of perceived discrimination and socioeconomic disadvantage that may discourage adequate engagement in healthcare and HIV/STI prevention services (Ayala, Bingham, Kim, Wheeler, & Millett, 2012). However, little is known about the mechanisms that link place to
acquisition of STIs other than HIV. Are these mechanisms the same as, or different from, the mechanisms that appear to link place to HIV infection?

In 2009, Earnshaw and Chaudoir proposed a theoretical model where three distinct stigma mechanisms, specifically internalized, anticipated, and enacted stigma, were linked to sexual and mental health outcomes. In their model, internalized stigma represents a sense of being less worthy or good or being inferior to others due to some characteristic about oneself (e.g., a health diagnosis). Internalized stigma has been shown in other studies to be directly linked to mental health outcomes, including depressive symptoms (Ayala et al., 2012). Anticipated stigma reflects an individual’s concern about how he or she expects to be treated (or mistreated) in the future by others due to some characteristic about him or herself. Medication adherence, attending healthcare visits, and other forms of engagement in healthcare may be most readily impacted by anticipated stigma (Kalichman et al., 2017). Finally, enacted stigma represents an individuals’ experiences of discrimination by others. Enacted stigma is most closely related to physical health concerns. Originally intended for the field of HIV, Earnshaw and Chaudoir’s stigma model arguably has utility for other STIs, such as gonorrhea, chlamydia, and syphilis infections, given that they also have significant stigma associated with them that can get in the way of optimal healthcare engagement (Ayala et al., 2012).

**Study Objectives**

The present study had three primary objectives. First, we aimed to describe the BMSM (i.e., the rates of HIV/STIs and internalized, anticipated, and enacted stigma they experience) and qualities about the areas in which they live (i.e., population density and poverty). Second, we aimed to determine associations between stigma, population density, poverty, and HIV/STI acquisition, in both bivariate and multivariable analyses. We hypothesized that there would be direct, positive
associations between stigma, population density, and poverty with HIV/STI acquisition. We also hypothesized that population density would moderate the relationships between stigma and poverty with HIV/STI acquisition.

Method

Sampling, Recruitment, and Enrollment

Participants (N=147) were recruited from gay-friendly bars, clubs, bathhouses, parks, and street locations in the greater Atlanta, GA metropolitan area, from online classifieds, and on social media (e.g., Facebook, Black Gay Chat, Jack’d). Participants were screened in-person using electronic handheld devices and over the phone using telephone screening software. For in-person screening procedures, recruiters approached participants as they entered the abovementioned target venues. Individuals were eligible to participate if they reported CAI in the past year with a man, tested HIV negative at the baseline appointment (using the OraQuick ADVANCE Rapid HIV-1/2 Antibody Test), and were at least 18 years of age. Individuals who tested HIV positive at the baseline appointment were linked to care and were referred to other studies for which they were eligible. Participants provided written informed consent for all study procedures, which included attending four in-person appointments at the study research site over a 12-month period. Each appointment included completing an assessment using Audio Computer Assisted Interviewing (ACASI) software, as well as oral, rectal, and urine STI testing for chlamydia and gonorrhea. A subsample was also tested for syphilis. Participants were compensated $30 for completing each study appointment, for a total of up to $120.

Measures

Social- and Structural-level Variables
While many studies of social- and structural-level factors on health behavior to date have used Zip Codes (Ransome Kawachi Braunstein & Nash, 2016), the present study instead used Census tracts. Census tracts are areas that are approximately equal to a neighborhood. Established by the Bureau of Census for analyzing populations, tracts generally encompass populations between 2,500 and 8,000 individuals. Census tracts are smaller than counties or cities, but larger than Census Block Groups or Blocks. The Census tract, rather than Census Block Group or Block, was the unit of analysis chosen due to the fact that the variables that are predicted to have an impact on the sample’s HIV/STI outcomes and HIV prevention health care behaviors are available at the tract level. All Census tract-level data used in the following analyses were retrieved from the 2010 U.S. Census.

**Neighborhood poverty.** Following published methods (Ransome et al., 2016; An, Prejean, McDavid, Harrison, & Fang, 2013) a principal component analysis was conducted to create community neighborhood poverty index scores by using Census tract-level poverty indicators based on data from the 2010 Census to characterize the level of poverty for every participant’s Census tract in our sample. We identified four indicator variables that have previously been shown to be associated with STI risk (Sonnenberg et al., 2013; Sullivan et al., 2011) and represent four domains of poverty: education, income, housing, and employment. “Percentage of residents with less than a high school education or GED” was the measure of education; “percentage of residents living at or below the federal poverty level” was the measure of income; “percentage of residents spending 30% or more of their incomes on housing” was the measure of housing; and “percentage of residents older than 16 years of age who are unemployed” was the measure of employment. The first component (herein referred to as the Neighborhood Poverty Index) accounted for 73% of the total variance and had a Cronbach’s α of 0.88. The index is a weighted linear combination of the
original four variables (correlation of each indicator with the index in parentheses): percentage of residents with less than a high school education or GED (0.38), percentage of residents living below the federal poverty level (0.35), percentage of residents spending 30% or more of their incomes on housing (0.39), and percentage of residents older than 16 years of age that are unemployed (0.39). We calculated the Neighborhood Poverty Index for the 356 Census tracts, and higher values reflected greater neighborhood poverty.

**Population density.** Population density of each participants’ Census tract was determined using the Census 2010 data. Population density by Census tract is equivalent to the 2010 population per square mile by Census tract.

**Individual-Level Variables**

Data on all individual-level variables, including measures of demographic information, enacted stigma, and STI test results, were obtained from the Think Twice study. Refer to Study 1 for additional details.

**Dependent variables.**

*HIV/STI infections.* Participants took lab-confirmed tests for gonorrhea, chlamydia, and syphilis at each of the four time points (i.e, baseline and the 3-month, 6-month, and 12-month follow-up appointments). Participants who tested positive at any appointment were linked to care to begin treatment for their infection. Therefore, participants who tested positive for any STI at more than one time point were considered to have separate infections. Finally, participants took an HIV test at the 12-month appointment.

**Independent variables.**

*Demographics.* Information on participants’ age, years of education, employment status, income level, race/ethnicity, and sexual orientation (i.e. whether they identified as same gender
loving/gay, bisexual, or heterosexual) at the baseline appointment of the Think Twice study (Eaton et al., 2018).

**Enacted stigma from health care providers.** In order to determine participants’ perceptions of enacted stigma from health care providers, they were asked six questions. The first three questions focused on their perception of discrimination due to their race, and included items such as “In the past year, do you think you have been mistreated by healthcare providers because of your race?” and “In the past year, do you think your healthcare isn’t as good as others because of your race?” The last three questions focused on participants’ perception of discrimination due to their sexual orientation, and included items such as “In the past year, do you think you have been ignored by healthcare providers because of your sexual orientation?” and “In the past year, do you think your healthcare isn’t as good as others because of your sexual orientation?” Participants were asked to answer either yes or no to each question, and participants were scored with a 1 for a yes response and a 0 for a no response. These items were adapted from a study by Wilson and Yoshikawa (2007). Given that race-based and sexual orientation-based discrimination in the healthcare setting were found to be highly collinear ($r = .72$), participants’ scores were summed across all six items to create one perceived discrimination variable (Cronbach’s $\alpha = .90$). These variables were also looked at separately and controlled for in each analysis.

**Statistical Analysis**

All data analyses were performed using PAWS Statistics, version 24.0 (SPSS, Inc., Chicago, IL), as well as the PROCESS macro for SPSS (Hayes, 2012). For any variable included in the analyses, less than 5% of the data were missing. In the moderation analyses, estimates were based on 10,000 bootstrapping replicates.
Descriptive statistics were conducted to obtain frequencies and percentages for the categorical variables, as well as means, standard deviations, and ranges for the continuous variables. For inferential statistics, both bivariate and multivariable analyses were performed using the following variables: neighborhood socioeconomic deprivation, rurality, perceived healthcare-related stigma, and STI diagnosis. Pearson correlations and t-tests were performed to identify significant associations among the abovementioned variables of interest. Moderation analyses were performed to identify the nature of the relationships between the independent variables (i.e., stigma, neighborhood poverty) and dependent variables (i.e., HIV/STI diagnosis).

To examine whether either population density moderated the relationships between (a) neighborhood poverty and (b) enacted stigma from health care providers with HIV and STI diagnoses, two regression models were constructed (henceforth referred to as Model 1 and Model 2). In Model 1, the following interaction terms were examined: neighborhood poverty × population density and stigma × population density with HIV status as the outcome variable. The same two interaction terms were examined in Model 2 with STI status as the outcome variable.

For any statistically significant (at the $p < 0.05$ level) moderation effect found, the Johnson-Neyman (J-N) procedure was employed (Preacher, Curran, & Bauer, 2006). The J-N procedure provides information about the percent of cases in the data with values of the moderator above or below the points of transition in significance. In sum, the J-N procedure was used to determine at what value of the moderator the independent variable would be statistically significantly related to the outcome variable (i.e., the “threshold”).

Finally, we calculated the population attributable fraction (PAF) for enacted stigma from health care providers with the overall rate of incident infections (HIV and other STIs). The population attributable fraction is defined as the proportional reduction in average disease risk
(e.g., HIV risk) that would be achieved by eliminating the exposure(s) of interest from the population while distributions of other risk factors in the population remain unchanged” (Rockhill, Newman, & Weinberg, 1998).

### Results

**Demographic Variables**

A total of 147 BMSM residing in the southeastern United States were included in the present analysis. Participants were, on average, 30.6 years of age ($SD = 10.3$). Of the total sample, 101 (68.7%) identified as gay, homosexual, or same gender loving; 44 (29.9%) identified as bisexual, and the remaining two (1.4%) identified as straight or heterosexual. Seventy-two (49.0%) BMSM reported annual incomes equal to or less than $10,000, and 31 (21.1%) reported having received a college degree (Table 1).

**Enacted Stigma, Neighborhood Poverty, and Population Density**

With respect to enacted stigma, 46 (31.2%) participants reported having experienced racial ethnic identity-based or sexual orientation-based healthcare-related stigmatization. Further, with respect to neighborhood poverty, scores ranged from -1.459 to 3.328, with higher scores indicating greater neighborhood poverty. The average neighborhood poverty score was 0.879, which reflected a moderate level of neighborhood poverty. Finally, of the Census tracts represented by participants’ places of residence, population density values ranged from a minimum of 27 persons per square mile to a maximum of 10,836 persons per square mile. The average population density was 4,573 persons per square mile, which, according to the U.S. Census, reflect geographic areas that are more urban or “city-like” in nature.

**Regression Analyses with Enacted Stigma, Poverty, Population Density, and HIV/STI Diagnoses**
The results of the present study indicate that population density and neighborhood poverty were significantly, positively correlated \((r = .31)\), suggesting that, at the Census tract level, as population density increased (i.e., as an idea became more urban or “city-like”), neighborhood poverty did as well. Further, BMSM who had experienced race- or sexual orientation-based stigma lived in significantly more impoverished areas \((t = 2.63, df = 146, p < .001)\). Finally, BMSM who lived in less densely populated Census tracts were more likely to test positive for HIV \((t = 2.83, df = 146, p < .01)\), but the relationship was not significant for other STIs.

**Main Effects in the Logistic Regression Models**

In Model 1, both neighborhood poverty \((\beta = 2.84, SE = 1.44, p = 0.05, CI: 5.66, 0.02)\) and population density \((\beta = 0.07, SE = 0.03, p = 0.02, CI: 0.13, 0.01)\) were significantly associated with testing positive for HIV, however, enacted stigma was not associated with testing positive for HIV.

In Model 2, none of the proposed variables – neighborhood poverty, population density, and enacted stigma – were significantly associated with testing positive for other STIs (Table 2).

**Interaction Effects in the Logistic Regression Models**

Population density significantly moderated the relationship between neighborhood poverty and testing positive for HIV \((\beta = 0.320, SE = 0.171, p = 0.05, CI: 0.641, 0.011)\). The estimate indicated that the significant, positive relationship of neighborhood poverty with testing positive for HIV became stronger by 0.002 for each one-unit increase in population density (i.e., for each one person increase to the population per square mile). This finding was in line with our original hypothesis that population density moderated the relationship between neighborhood poverty and testing positive for HIV; however, there was no evidence to support our other hypotheses.
In order to understand how the relationship between neighborhood poverty and testing positive for HIV varies at different population density values, simple (conditional) slopes were estimated for the association between neighborhood poverty HIV status at different levels of population density (sample range = 27-10,836). The values of population density of 1,488 persons per square mile, 4,573 persons per square mile, and 7,658 persons per square mile, which represent the mean – 1 SD, the mean of the sample, and the mean + 1 SD, were chosen. The simple slopes were estimated, and they were found to be statistically significant at 4,573 persons per square mile and at 7,658 persons per square mile, but not at 1,488 persons per square mile. The simple slope value at 7,658 persons per square mile (1.181, \( p < .05 \)) was more positive than the value at 4,573 persons per square mile (0.740, \( p < .05 \)), meaning that the simple slope became more positive as population density increased. Otherwise stated, neighborhood poverty was significantly, positively associated with testing positive for HIV for persons living in more urban areas with population densities of 4,573 persons per square mile and 7,658 persons per square mile, but not at 1,488 persons per square mile, and that this relationship became stronger as population density increased. The J-N procedure was used to find the range of population densities for which the conditional slopes between the independent variable (i.e., neighborhood poverty) and the dependent variable (i.e., HIV status) is statistically significant. It was found that the relationship between neighborhood poverty and HIV status became statistically significant at, and remained significant for all Census tracts with population densities above, 2,180 persons per square mile (Table 2).

**Population Attributable Fraction: Enacted Stigma**

Population attributable fraction (PAF) was calculated to determine the attributable risk for HIV and other STIs that enacted stigma (both race- and sexual orientation-based stigmas) from
health care providers carried; in other words, PAF was calculated to determine what proportion of HIV/STI cases would be averted if not a single case of enacted stigma (i.e., discriminatory experiences due to one’s race or sexual orientation) had occurred. PAF calculations revealed that the attributable risk percent for enacted stigma on HIV infection was 10.127% and for other STIs was 5.431%, meaning that 10.127% of HIV cases and 5.431% of other STI cases may have been averted in this sample had discriminatory behaviors and/ practices not occurred against them (Table 3).

Discussion

Results from the present study indicate that a number of variables beyond the individual-level are relevant with respect to HIV infection, namely neighborhood poverty and population density. The finding that both variables were significantly, positively associated with HIV infection was in line with our proposed hypothesis regarding these variables, as well as prior literature which suggested that neighborhood poverty’s association with HIV infection would be strongest in more urban areas (Vaughan et al., 2014). While there were no significant findings in the logistic regression models related to enacted stigma, nor were any of the variables associated with STIs, these findings offer some guidance for future research efforts in that they suggest (1) attending to social/structural risk variables and (2) acknowledging that working with urban and rural populations will require attending to social/structural risk variables that are unique to those environs.

The present study’s findings may best be interpreted using both Poundstone et al.’s (2004) social epidemiological model, in addition to a similar framework, referred to as the syndemics model. While the syndemics model was not originally used to conceptualize this study in the way that the social epidemiological model was, the syndemics model may offer additional perspective
on how best to interpret the present study’s findings. According to Singer and his colleagues (2017, p. 941), the syndemics model of health focuses on the “biosocial complex, which consists of interacting, co-present, or sequential diseases and the social and environmental factors that promote and enhance the negative effects of disease interaction.” The authors posit that, conventionally, diseases have been understood as distinct entities that separate from other diseases and independent of the social and structural contexts in which they occur. A syndemics approach, however, provides an alternative orientation to disease management – and health promotion – by demonstrating how an *integrated* approach to studying and addressing diseases may be especially useful and successful. This approach, which involves recognizing the co-occurrence of social/structural and health issues, shows how non-biomedical/pharmacological interventions are relevant and have the potential to substantially change key health outcomes. By requiring a “big picture” awareness of disease, disease clustering, and disease interactions in ecological contexts, rather than considering disease from a traditional, biomedical, narrower scope, this framework pushes us to consider approaches to disease prevention and management, as well as health promotion, that involve not only behavioral interventions, but also changes to social polices around issues like poverty (Singer & Clair, 2003).

The present study supports, and argues for further expansion of, the syndemics model. First, in our study we found that there were contextual factors beyond the individual-level that were related to HIV infection, namely neighborhood poverty and neighborhood population density. These findings support the notion that social and environmental factors *do* facilitate disease, and that disease is not an entity independent from social/structural contexts. Second, the results of the present study expand beyond the syndemics model, in that they offer evidence that the social and environmental facilitators of disease vary across contexts (Heckman et al., 1998).
Our finding that neighborhood poverty was significantly associated with HIV infection in urban areas, but not in more rural areas, suggests that there may be something different about neighborhoods that have higher levels of neighborhood poverty – or individuals who live in neighborhoods with higher levels of neighborhood poverty – in urban areas as compared to those in rural areas. Previous studies which have explored income inequality as a determinant of health have argued that income inequality is (1) greater in urban environments than in rural environments, and that (2) greater income inequality is associated with a variety of troublesome health outcomes (Florida & Mellander, 2016). It is possible that neighborhood poverty was associated with HIV infection in the present study in urban and not rural environments due to greater income inequality in the general area (i.e., in surrounding Census tracts), however, as income inequality was not measured in this study, additional studies are warranted to investigate this hypothesis. Prior studies (Holtgrave & Cosby, 2003; Holmqvist, 2009) have argued that income inequality promotes disinvestment in human capital, i.e., less trust amongst citizens, less investment in community services, and the non-assumption of reciprocity between citizens. This disinvestment in human capital often accompanies other issues, like neighborhood social disorder (e.g., crime, delinquency, drug use) and stress and other mental health concerns due to neighborhood disorder (e.g., depression, anxiety; Latkin, Williams, Wang, & Curry, 2005; Latkin & Curry, 2003). Kerrigan and his colleagues (2006) argued that areas of concentrated disadvantage – essentially, high rates of all of the above: poverty, crime, drug use and trade, and mental health concerns – are especially risky with respect to promoting HIV risk behaviors. It is potentially the case that urban environments have higher rates of concentrated disadvantage, including neighborhood poverty, than do rural areas that are impoverished. Additional studies are warranted, and may consider approaching these questions using the syndemics model (Singer & Clair, 2003). While the
syndemics model has been predominately used to consider the clustering of health conditions, it may also be used to understand the dynamic interplay and clustering of social conditions which facilitate risk and acquisition of HIV.

While enacted stigma from health care providers was not significantly associated with HIV or STI, we did find interesting attributable risk results from our PAF calculations, namely that approximately 10% of people with HIV and 5% of people with STI could have had their infections averted without the presence of stigmatizing events due to race or sexual orientation from health care providers. This finding provides preliminary evidence that race- and sexual orientation-based stigma from health care providers undermines our efforts to prevent HIV in key populations (e.g., BMSM), and that eliminating these forms of stigma may be key in promoting open, healthy patient-provider relationships and in reaching our HIV prevention goals. Potential avenues through which race- and sexual orientation-based stigma from health care providers increases risk for HIV/STI include potentially dissuading individuals from regularly attending care, which may lessen the likelihood of knowing one’s status, accessing biomedical HIV prevention options, and increase psychological distress. Educational programs which include modeling of non-stigmatizing behavior may be necessary to teach health care providers how to provide good quality, non-stigmatizing care and to establish trusting relationships with patients who are racial and/or sexual orientation minorities. It is worth noting, however, that the present study’s results are based on a relatively small sample size, and so additional research with bigger sample sizes is needed to better understand enacted stigma from health care providers among BMSM, including whether and to what extent intersectional stigma (i.e., a dynamic interaction between multiple stigmatized identities) is experienced.

Limitations
The findings of the current study should be interpreted in light of their limitations. The current sample is isolated to one region of the country; therefore, these findings may not generalize to the larger population of BMSM in other areas of the U.S. Further, the sample size is relatively small, and data for this study are cross-sectional were collected using self-report measures, which are prone to social desirability bias. Additionally, measures of race- and sexual orientation-based enacted stigma were limited in that they each only included a single, dichotomous item, which were merged due to being highly correlated. More comprehensive measures, including measures of intersectional race and sexual orientation stigma, are needed for future analyses.

**Conclusions**

We found that neighborhood poverty and population density interacted, such that neighborhood poverty in more densely populated areas, and not in more rural areas, was associated with HIV infection. These findings indicate that not only are factors beyond the individual-level important, but those factors vary in their importance by geographic context. In the future, interventionists must consider the extent to which the environment in which their participants reside is urban or rural, and consider social and structural facilitators of risk and disease that are unique to those environs. In addition, the study suggests that simply considering low income as a determinant of risk, devoid of geographic context, is important but not sufficient: low income in an urban area may be characteristically different from low income in a rural area. Additional research is warranted to consider whether urban, low-income areas are more prone to concentrated disadvantage than are rural, low income areas, and to consider whether neighborhood poverty in urban areas is associated with other forms of social and structural disadvantage that impede our efforts toward preventing HIV.
References


study of the impact of neighborhood disorder. *Journal of Health and Social Behavior*, 34-44.


Table 1

Descriptive Statistics for the Sample of 147 Black Men Who Have Sex with Men (BMSM) Located in the Atlanta, GA Metropolitan Area

<table>
<thead>
<tr>
<th>Variable</th>
<th>M (range)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Census tract level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neighborhood Poverty</td>
<td>0.879 (-1.459</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>to 3.328)</td>
<td></td>
</tr>
<tr>
<td>Population Density</td>
<td>4,573 (27-</td>
<td>3,085</td>
</tr>
<tr>
<td></td>
<td>10,836)</td>
<td></td>
</tr>
<tr>
<td><strong>Individual level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>30.61 (18-59)</td>
<td>10.28</td>
</tr>
<tr>
<td>Education</td>
<td>1.92 (1-5)</td>
<td>1.03</td>
</tr>
<tr>
<td>Income</td>
<td>2.27 (1-7)</td>
<td>1.66</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sexual Orientation</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gay/homosexual/same gender loving</td>
<td>101</td>
<td>68.7</td>
</tr>
<tr>
<td>Bisexual</td>
<td>44</td>
<td>29.9</td>
</tr>
<tr>
<td>Straight/heterosexual</td>
<td>2</td>
<td>1.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race- and Sexual Orientation-Based Healthcare-Related Enacted Stigma</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, both race- and sexual-orientation-based discrimination</td>
<td>25</td>
<td>17.0</td>
</tr>
<tr>
<td>Yes, race-based discrimination only</td>
<td>10</td>
<td>6.8</td>
</tr>
<tr>
<td>Yes, sexual orientation-based discrimination only</td>
<td>11</td>
<td>7.5</td>
</tr>
<tr>
<td>Neither race- nor sexual orientation-based discrimination</td>
<td>101</td>
<td>68.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HIV/STI Infections</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>Gonorrhea</td>
<td>47</td>
<td>-</td>
</tr>
<tr>
<td>Chlamydia</td>
<td>45</td>
<td>-</td>
</tr>
<tr>
<td>Syphilis</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note.* Men who identified as straight/heterosexual were included in the analysis if they reported at least one instance of engaging in male-to-male sexual contact in the last year.
Table 2

*Results from the Fixed Effects Regression Analyses Showing the Moderation Effect of Population Density on the Relationship Between Neighborhood Poverty and HIV Infection*

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimate</th>
<th>SE</th>
<th>p</th>
<th>CI (lower)</th>
<th>CI (upper)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.48</td>
<td>0.03</td>
<td>0.67</td>
<td>-0.01</td>
<td>0.76</td>
</tr>
<tr>
<td>Neighborhood poverty → HIV</td>
<td>2.84</td>
<td>1.44</td>
<td>0.05</td>
<td>0.02</td>
<td>5.66</td>
</tr>
<tr>
<td>Enacted stigma → HIV</td>
<td>0.002</td>
<td>0.001</td>
<td>0.78</td>
<td>-0.28</td>
<td>0.89</td>
</tr>
<tr>
<td>Population density → HIV</td>
<td>0.07</td>
<td>0.03</td>
<td>0.02</td>
<td>0.13</td>
<td>0.21</td>
</tr>
<tr>
<td>Poverty × Density → HIV</td>
<td>0.32</td>
<td>0.17</td>
<td>0.05</td>
<td>0.11</td>
<td>0.64</td>
</tr>
<tr>
<td><strong>Model 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.29</td>
<td>0.11</td>
<td>0.45</td>
<td>-0.11</td>
<td>5.67</td>
</tr>
<tr>
<td>Neighborhood poverty → STI</td>
<td>0.89</td>
<td>0.45</td>
<td>0.31</td>
<td>-0.45</td>
<td>1.16</td>
</tr>
<tr>
<td>Enacted stigma → STI</td>
<td>0.07</td>
<td>0.01</td>
<td>0.87</td>
<td>-0.11</td>
<td>0.02</td>
</tr>
<tr>
<td>Population density → STI</td>
<td>0.11</td>
<td>0.08</td>
<td>0.43</td>
<td>-0.43</td>
<td>0.62</td>
</tr>
</tbody>
</table>

*Note.* For each of the two models, the 95% CIs were obtained by bias-corrected bootstrapping with 10,000 resamples. In Model 1, neighborhood poverty and enacted stigma are the independent variables \((X_1, X_2)\), population density is the moderator variable \((M)\), and HIV status is the outcome variable \((Y)\). In Model 2, neighborhood poverty and enacted stigma are the independent variables \((X_1, X_2)\), population density is the moderator variable \((M)\), and STI status is the outcome variable \((Y)\). CI (lower) = lower bound of a 95% confidence interval; CI (upper) = upper bound of a 95% confidence interval.
Table 3

Results from the PAF Calculations for Enacted Stigma and HIV and STI Cases

<table>
<thead>
<tr>
<th>Variable</th>
<th>PAF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enacted stigma → HIV cases</td>
<td>10.127</td>
</tr>
<tr>
<td>Enacted stigma → STI cases</td>
<td>5.431</td>
</tr>
</tbody>
</table>
Using Twitter for remote monitoring of public attitudes toward, and activity around, tenofovir disoproxil fumarate (TDF)/emtricitabine (FTC) as pre-exposure prophylaxis (PrEP)
Chapter 4: Using Twitter for remote monitoring of public attitudes toward, and activity around, tenofovir disoproxil fumarate (TDF)/emtricitabine (FTC) as pre-exposure prophylaxis (PrEP)

Social media produces massive amounts of data on a large scale. Twitter, Facebook, Instagram, and other widely utilized social media outlets encourage frequent user expressions of their thoughts, opinions, and other details of their lives. According to a study by Pew Research Center (Smith & Brenner, 2012), approximately 15 percent of all online adults use Twitter. Twitter is a popular social media site that allows its users to post publicly available, real-time communications or “tweets.”

Researchers at the Pew Research Center (Smith & Brenner, 2012) surveyed U.S. Twitter users and found that they are, compared to the general U.S. population, more likely to be African American/Black, not heterosexual, and younger in age (Pew Research Center, 2012). Because there is a growing HIV epidemic among African-American/Black gay, bisexual, and other young men who have sex with men (MSM) Twitter may make it a particularly useful resource for the intersection between the field of HIV and the new and emerging field of “digital epidemiology,” which studies how “big data” can be used to understand, detect, and address public health issues (Young, Rivers, & Lewis, 2014).

Previously, researchers have shown that analyzing tweets can measure a variety of population characteristics, including public opinion and public health measures. Tweet content has been used to predict events, such as election results (Tumasjan et al. 2010; O’Connor et al. 2010), as well as influenza outbreaks and mentions of a variety of common ailments, including allergies, obesity, and insomnia (Culotta, 2010; Paul & Dredze, 2011). Few studies have ventured into social and behavioral HIV prevention science, and, therefore, Twitter remains a relatively untapped
resource in this area. However, one study by Young and his colleagues (2014) found Twitter to be an effective tool for evaluating and detecting HIV risk behaviors and outcomes.

In other health fields, tweets have been analyzed for mentions of a variety of common ailments, including allergies, cancer, obesity, and insomnia (Lee, Agrawal, & Choudhary, 2013; Paul & Dredze, 2011). Twitter has also been used to track illnesses over time (i.e., syndromic surveillance), measure behavioral risk factors, localize illnesses by geographic region, and analyze symptoms and medication usage (Paul & Dredze, 2011), particularly in the context of influenza. A study by Lee and her colleagues (2013) showed that tweets that are “geotagged” (i.e., tweets where users have allowed for their location (via latitude and longitude) to be seen by other users) can be used to monitor and predict where outbreaks of the flu occur. In their study, Lee et al. mined six million flu-related tweets for keywords relating to common flu symptoms (e.g., headache, nausea) and medication usage (e.g., Tamiflu).

Despite these advancements in using “big data” from Twitter to better understand individuals’ behavior around influenza, there have been very few studies that use Twitter to survey the public’s attitudes toward or behaviors relating to HIV or HIV-related factors. This is particularly surprising given how well-versed many HIV researchers are in harnessing the power of social networking technologies for participant recruitment (Sullivan et al., 2011) and to serve as “venues” for behavioral interventions (Bull et al., 2012; Young et al., 2013). Additionally, given that young MSM of color are a key target population for HIV prevention efforts due to their disproportionately high risk for infection, understanding sentiment about various HIV prevention tools on Twitter is a logical step. According to a Pew Research Center study (Smith & Brenner, 2012), the site has a high proportion of users who identify as African-American/Black and young. The study found that 14% of adults aged 18 to 29 use Twitter, compared to less than 4% of
individuals aged 65 and older, and 13% of African-American or Black individuals use Twitter compared to 5% of White individuals.

Despite its overwhelming potential, there has been limited use of Twitter to understand individuals’ attitudes toward HIV prevention technologies like PrEP. Once daily pre-exposure prophylaxis (PrEP) containing emtricitabine and tenofovir disoproxil fumarate (FTC-TDF) represents an efficacious and effective strategy for reducing the number of incident HIV infections among high risk populations, including men who have sex with men (MSM) and injection drug users (IDU). In the iPrEx and Partners PrEP trials, the incidence of HIV infection was reduced by 44% and 67%, respectively (Baeten et al., 2012; Grant et al., 2010).

Furthermore, in the context of HIV prevention, Twitter may be used as a geographic surveillance tool to help researchers make epidemiological inferences about rates of future or recent past behaviors, e.g., HIV risk behaviors, or to glean location-specific discussions around HIV and/or HIV medications. The resulting insights may help to facilitate faster and more targeted responses to, and preparation for, health epidemics, including those relating to HIV. Additionally, these insights may be helpful to healthcare professionals, community-based organizations (CBOs), and policymakers, in that they have the potential to aid in decision-making around targeted treatment efforts, marking campaigns, and allocation of other limited resources, including disease prevention technologies (e.g., free HIV testing services, pre-exposure prophylaxis (PrEP) marketing campaigns).

It is worth noting that relatively little research has been conducted that examines social and structural factors and PrEP attitudes. Most PrEP attitudes research has considered the direct impact of peers or friends (Smith, Toledo, Smith, Adams, & Rothenberg, 2012) or healthcare providers (Tellalian, Maznavi, Bredeek, & Hardy, 2013). However, to date, no study has considered whether
social and structural characteristics of an individual’s place of residence (e.g., HIV prevalence, health insurance coverage) are associated with attitudes toward PrEP.

**Study Objectives**

The present study had three primary objectives. First, we determined whether and to what extent tweet frequency (about Truvada as PrEP) was related to users’ locations. We hypothesized that tweet frequency would be significantly higher in Metro Areas where PrEP roll-out programs and marketing campaigns have been greatest (i.e., San Francisco, CA, New York City, NY). Second, we sought to determine whether and to what extent sentiment (about Truvada as PrEP) in the tweets was related to users’ locations. Again, we hypothesized that sentiment would be significantly more positive in Metro Areas where PrEP roll-out programs and marketing campaigns have been greatest. Third, we evaluated whether three area-level variables (i.e., HIV prevalence, health insurance coverage) were associated with (a) tweet frequency about Truvada as PrEP and (b) sentiment about Truvada as PrEP in tweets. We hypothesized that there would be positive associations between HIV prevalence and health insurance coverage with tweet frequency and positive sentiment about Truvada as PrEP.

**Method**

This study received exemption from the University of Connecticut Institutional Review Board (IRB). Tweets \(N = 49,354\) including one or more of the non-case sensitive key words (“Truvada” and “#PrEP) were collected from Twitter’s free Application Programming Interface (API) between June 15\(^{th}\), 2017 and December 15\(^{th}\), 2017.

When an individual makes a search on Twitter on his or her web browser (e.g., Safari, Google Chrome, et cetera), the individual’s computer sends a request to Twitter. In return, Twitter sends back a web page representation of the requested tweets. The Twitter API functions
similarly in that similar requests can be made by compute programs to have tweets returned in a format that is easy for the computer program to use. The Twitter API’s responses to computer programs’ requests contain tweet text and several attributes about the tweets (called “metadata”). These data are able to be stored in a particular file called a JSON file, where tweets can later be mined for research purposes.

Due to the Twitter API’s case-insensitive keyword search, a proportion of the tweets in our dataset which included the phrase “#prep” were spurious. As an example, the following tweet was picked up by the Twitter API due to its inclusion of the phrase “#prep”; however, it references meal preparation and not Truvada as PrEP: “My meal prep game is fiercer than yours #mealprep #meal #prep #lunch #lunches”.

To address this issue, we removed tweets that were not related to Truvada as PrEP from the dataset (such as the example tweet above) by using code written in the Python programming language. This code allowed us to detect and retain the tweets that were related to PrEP (case sensitive) and remove the tweets that did not. Relevant Tweets were identified by including tweets that included the non-case sensitive word “Truvada” and the case sensitive word “#PrEP.”

To ensure that tweets including case sensitive phrases such as “#prep” or “#Prep” were not relevant, we used Python programming language to include tweets that had relevant key words in them, including: HIV, gay, MSM, and medication.

A sub-set of one percent of the retained tweets, as well as a sub-set of one percent of the rejected tweets, were evaluated to determine the reliability of our code at determining tweets that were relevant and tweets that were not relevant. The code was 99% reliable at determining relevant (retained) tweets, and 96.7% reliable at determining not relevant (rejected) tweets.
Upon completion of the above-mentioned procedures where we excluded the irrelevant tweets and included the relevant tweets, we retained a total of \( n = 33,936 \) tweets for our analyses, which are described in further detail in the “Text Analysis” and “Geographical Analysis” subsections below.

**Tweet Metadata**

The Twitter API is able to provide a variety of metadata for each tweet, in addition to the tweet’s content, including the user’s (i.e., the person who sent the tweet) language, his or her number of followers (i.e., people who subscribe to the user’s communications), his or her location, the number of “likes” and “re-tweets” that each tweet receives (i.e., the tweet’s “endorsements” from other Twitter users), and the date and time that the tweet was sent. Additionally, users may choose to enable a feature that links his or her location, in the form of latitude and longitude coordinates, to his or her tweets. If users enable geolocated data, then this information is provided through the API, and this information may be different from the location that is reported in the profile. For example, a person may tweet from a location not listed in their profile (e.g., if they are on a vacation). Additionally, geotagged tweets offer a much more precise location (i.e., latitude and longitude coordinates) than do profile locations, which are usually described more generally, e.g., “Washington, D.C.,” “NorCal,” or “New York State.”

While geotagged tweets are preferable to simply relying on users’ location descriptions in their Twitter profiles, relatively few \( n = 894 \) users in the present sample opted-in to enable geolocated data for their tweets. As a result, for all tweets that were not geotagged, the authors’ locations, which were described in their profiles, were utilized to infer those users’ probable locations. For example, for a user who does not have geotagged tweets but has “Washington, D.C.” listed in her profile, the Google API can be used to assign the user a set of latitude and longitude
coordinates of the centroid of a specified area (in this case, the user would be assigned the latitude and longitude coordinates of the center of the city of Washington, D.C.).

By assigning a probable location to users without geotagged tweets but with a valid location description in their profile, the number of tweets with an assigned (known or estimated) location in our dataset increased substantially. The remaining users not included in the geographic analyses did not have geotagged tweets and either did not list a location in their profile or listed an invalid location (e.g., “Neverland”). Therefore, users with geolocated tweets, as well as users with self-reported, valid locations in their profiles, were included in all location-based analyses for the present study.

Only geolocated tweets from users located in the United States MSAs of interest were used, which limited the dataset to \(n=7,810\) for the geographic analyses. Tweets were limited to only those in the United States because we were interested to know whether sentiment of the tweets was related to the HIV prevalence and social determinants of health in the area (e.g., number of people without health insurance) at the county level.

**Text Analysis**

The overarching goal of text analysis of tweets in the context of using Twitter data is to reveal deep insights by examining the content of those tweets. Broadly, text analysis of tweets is particularly useful because Twitter, as compared to more traditional online communication tools (such as blogs, forums, mailing lists, et cetera), has many users, i.e., approximately 330 million. So many individuals utilize Twitter in part due to its ease of accessibility and popularity, i.e., potential for greater interaction with other users (Pak & Paroubek, 2010). With greater numbers of individuals discussing their thoughts and opinions about products and services that they use, their
political views, and other potentially important information, Twitter is a valuable resource for data, as it serves as a hub for many people’s thoughts and sentiments about a variety of topics.

In the case of the present paper’s text analysis, IBM Watson® Natural Language Understanding (NLU) API is used. Watson® NLU API is a cloud-based service with a variety of features that may be used to both extract and analyze information from text, including tweet text. For the purposes of the present paper, the following Watson® NLU features are used: Sentiment, Emotion, Concepts, and Keywords, all of which are described in detail below.

**Sentiment.** Twitter sentiment analysis involves determining the sentiment of tweets, or an estimate of the user’s personal positive, neutral, or negative feelings contained in a tweet’s text. In general, Twitter sentiment analysis may be particularly useful for researchers, marketers, and policymakers with a vested interest in Truvada as PrEP, particularly those who wish to better understand public opinion of the drug.

The Watson® NLU API is a strong choice for sentiment analysis in particular for this paper, given that the text content that will be analyzed is from tweets, and the NLU API sentiment analysis engine was taught how to determine a text’s sentiment based on data specifically from Twitter. This means that the NLU API can handle emoticons and informal language quite well as compared to other APIs.

In addition, the NLU API runs on a *recurrent neural network*, which is currently the premier tool in the natural language processing arena. While some APIs with sentiment analysis capabilities use pre-established “dictionaries” of positive-, negative-, and neutral-assigned words (and then simply detect those words in the text input to determine its overall sentiment), recurrent neural networks are capable of understanding long-range dependencies within a string of text.
Specifically, the NLU API is able to understand sentence structure, as opposed to simply detecting presence or absence of positive, neutral, and/or negative words.

The NLU API assigns each tweet one of three categories: “negative,” “neutral,” or “positive.” In addition, the NLU API assigns each tweet a numeric value ranging from -1 to 1, where smaller values represent more negative sentiment, and larger values represent more positive sentiment. For the present paper, information about tweets’ sentiment is collected in the form of categories and the numeric value ranging from -1 to 1. Finally, the NLU API is able to determine sentiment not just about the overall text it is provided, but also towards specific, user-specified keywords in the text, like PrEP, Truvada, healthcare, LGBT rights, and so on. This is useful because a tweet like “I want to be on PrEP. I’m mad my doctor doesn’t know anything about it” might be classified by the NLU API as negative (due to the author’s upset over his or her doctor’s lack of knowledge about PrEP), however, the author’s sentiment toward PrEP is positive. Therefore, for the present paper, sentiment was determined in reference to the keywords PrEP and Truvada to ensure that conclusions drawn about sentiment were specifically in reference to the medication.

**Emotion.** With the Watson® NLU API, the following emotions can be detected in text: joy, sadness, anger, disgust, and fear. Emotion information, like sentiment information, can be returned for the overall text or for user-specified target phrases or keywords. Each tweet receives a value for each of the five emotions (i.e., joy, sadness, anger, disgust, and fear) ranging from 0 to 1, where higher scores reflect higher levels of that emotion.

**Geographical Analysis**

For the geographical analyses, we focused on the location-based tweets only, which included tweets with geolocated data as well as those from users with profile location descriptions.
Data were filtered to include only those tweets that had a location in the United States Metropolitan Statistical Areas (MSAs) represented in AIDSVu, limiting the sample to \( n=5,007 \) tweets.

Using the Google Geocoding API, we were able to retrieve a latitude/longitude coordinate location for all tweets whose authors had a profile location listed, but did not enable geolocation services for their tweet. This method is subtly different than using geolocation from the tweet metadata, since it reflects a location that each user self-identifies as his or her home, but it may not necessarily be his or her current location. A variable was encoded for each tweet indicating whether location was gotten from tweet metadata coordinates or from user profile home location geocoded by Google. However, all tweets with latitude/longitude coordinates, either from the tweet metadata or from the Google Geocoding API (which used the user’s profile location).

Due to the fact that the overwhelming majority of the tweets about Truvada as PrEP were from larger cities, the level of analysis used for geographic area was the Metropolitan Statistical Area (MSA). Therefore, all tweets within a Metropolitan Statistical Area were labeled according to as Federal Information Processing Standard (FIPS) codes. Metropolitan Statistical Areas are defined as core urban areas with populations equal to or greater than 50,000 persons that have a high degree of social and economic integration (for example, as measured by commuting to work). Example U.S. Metropolitan Statistical Areas include the Washington, D.C./Arlington, V.A./Alexandria, W.V. Metro Area, the Fresno, C.A. Metro Area, and the Atlanta/Sandy Springs/Roswell, G.A. Metro Area. For the purposes of the present paper’s analyses, tweet attributes, specifically were compared with Metro Areas attributes.

**Sentiment.** As described above, each tweet, including the location-based tweets, was assigned a sentiment category (i.e., negative, neutral, or positive) and numeric value ranging from
-1 to 1 (where higher values indicate more positive sentiment) where the sentiment was determined with respect to the keywords Truvada and #PrEP (rather than the sentiment of the overall tweet).

After all location-based tweets were assigned a sentiment category and numeric value, as well as a Zip Code and, if possible, a Metropolitan Statistical Area, ArcMap was used to visually represent public sentiment about Truvada as PrEP across locations in the U.S. Specifically, Zip Codes and Metropolitan Statistical Areas will be visually compared by creating colored maps of these areas, where blue represents more negative sentiment and red represents more positive sentiment.

**HIV prevalence.** HIV prevalence data were extracted from AIDSVu.org and assigned to Zip Codes and U.S. Metropolitan Statistical Areas represented by the location of the tweets.

**Health insurance coverage.** Percent of residents with health insurance was determined using most recent Census Bureau data from 2010.

**Statistical Analysis**

In addition to producing visual representations of the data (i.e., providing maps of tweets’ location), simple statistical analyses were performed to determine associations between the above-mentioned variables. Univariate regressions assessed associations between (a) number of tweets (i.e., frequency), (b) the sentiment of the tweets, (c) area HIV prevalence, and (d) health insurance coverage.

**Results**

**Sentiment**

All relevant tweets ($N=7,810$) were run through Watson® NLU API to determine an overall sentiment score for Truvada/PrEP for each tweet. The average sentiment score for the total sample was 0.059, indicating a slightly positive attitude overall towards Truvada as PrEP for HIV
prevention. The three MSAs with the most negative tweet sentiment score, on average, were Baton Rouge, LA ($M = 0.017$), Raleigh, NC ($M = 0.036$), and Atlanta, GA ($M = 0.064$), and the three MSAs with the most positive tweet sentiment score, on average, were San Francisco, CA ($M = 0.815$), New York City, New York ($M = 0.786$), and Washington, D.C. ($M = 0.758$).

**Emotion**

Watson NLU API was also used to determine emotion scores regarding fear, joy, and disgust towards Truvada/PrEP for the sample ($N=7,810$). The sample’s average fear emotion score was -0.230, average joy score was 0.191, and average disgust score was -0.348. These scores reflected that, among these three emotions, the sample had the lowest levels of feeling disgust towards PrEP, and the highest levels of feeling joy towards PrEP. The three MSAs with the highest levels of joy towards PrEP were San Francisco, CA ($M = 0.709$), San Diego, CA ($M = 0.669$), and New York City, NY ($M = 0.519$). The three MSAs with the highest levels of fear towards PrEP were Jackson, MI ($M = 0.551$), Raleigh, NC ($M = 0.451$), and Las Vegas, NV ($M = 0.359$). Finally, the three MSAs with the highest levels of disgust towards PrEP were Dallas, TX ($M = 0.451$), Birmingham, AL ($M = 0.451$), and Jackson, MI ($M = 0.451$).

**Fixed Effects Regression with HIV Prevalence Rate, Health Insurance Coverage, Tweet Frequency, and Tweet Sentiment**

The geographical analysis was limited to $n=5,007$ tweets located in the MSAs for which AIDSVu had HIV prevalence and health insurance coverage data.

First, a fixed effects Poisson regression with tweet frequency as the outcome variable and MSA as the grouping variable and a random intercept was conducted with HIV prevalence rate and health insurance coverage as independent variables. HIV prevalence rate ($\beta = 0.779$, $SE = 0.147$, $p = 0.042$), but not health insurance coverage, was significantly, positively associated with
tweet frequency. In other words, MSAs with higher HIV prevalence rates also had significantly more tweets about Truvada/PrEP.

Second, a fixed effects linear regression with tweet sentiment as the outcome variable and MSA as the grouping variable and a random intercept was conducted with HIV prevalence rate and health insurance coverage as the independent variables. Both HIV prevalence rate ($\beta = 0.206$, $SE = 0.111$, $p = 0.039$) and health insurance coverage ($\beta = 0.567$, $SE = 0.244$, $p = 0.010$) were significantly associated with tweet sentiment. Both HIV prevalence rate and health insurance coverage were significantly, positively associated with tweet sentiment, suggesting that MSAs with higher HIV prevalence rates and greater proportion of health insurance coverage amongst residents had more positive sentiment toward Truvada as PrEP.

**Discussion**

The results of the present study serve the first piece of evidence for how social media data from Twitter might be used for extracting information regarding individuals’ attitudes toward Truvada as PrEP for HIV prevention. Results suggest that using data from Twitter to detect Truvada sentiment is feasible, and that identifying the tweets from various MSAs of interest is also feasible. With the existence of a plethora of HIV surveillance data and social determinants of health data at various geographic levels (including the MSA level), linking Twitter sentiment information to relevant qualities of geographic areas is feasible, and it offers insights into how sentiment reflected in tweets might relate to qualities of the larger geographic context in which the author of the tweet may reside. The study revealed that tweet frequency and tweet sentiment were positively related to HIV prevalence rate, and that tweet sentiment was positively related to proportion of the population with health insurance coverage at the MSA level. This study is important because (a) it provides support for the use of “big data” to survey discussions about, and attitudes toward,
topics of interest (i.e., PrEP for HIV prevention), and (b) a potential avenue for which future researchers and public health departments can capitalize on the ever-growing amount of social media data to monitor and detect areas that may be most in need of intervention (Young, 2015; Ozkose, Ari, & Gencer, 2015).

Sentiment about Truvada as PrEP was most positive in MSAs with higher HIV prevalence rates and higher rates of health insurance coverage. Frequency of discussions about Truvada as PrEP was higher in MSAs with higher HIV prevalence rates. These findings indicate that discussion around PrEP is occurring on Twitter most frequently in areas with the greatest rates of HIV prevalence, which is a promising finding given that awareness of PrEP amongst HIV negative individuals who are at elevated risk for HIV must be high in areas with the highest HIV prevalence rates (and potential higher network-level risk). It is also promising that tweet sentiment is positive in areas with greater HIV prevalence rates. However, the positive findings must be considered also in light of the potential issues: MSAs with lower, but certainly not negligible rates of HIV prevalence, have more negative sentiment toward PrEP, and discuss PrEP on social media platforms less often. Taken together, the present study’s findings suggest that, in areas with higher HIV prevalence rates, conversation about Truvada as PrEP is more frequent and more positive. Tweet sentiment was most positive in cities such as San Francisco, CA, New York, NY, and Washington, D.C., which have moderate to high HIV prevalence rates, but also generally liberal attitudes where Truvada as PrEP may be more accepted and less stigmatized than in more conservative areas where Truvada use may be seen as promiscuity-promoting and “slutty” (Spieldenner, 2016).

On the other hand, the present study revealed that tweet sentiment was negatively related to health insurance coverage rate, such that areas with less of the population with health insurance
had more negative sentiment toward PrEP for HIV prevention. This finding may be explained in one of two possible ways. First, areas with less health insurance coverage may indicate the populace, on average, being less connected to medical services, which may increase risk of misinformation regarding Truvada as PrEP. Secondly, areas with less insurance coverage may, very reasonably, have citizens who are less enthusiastic about, and less willing to take, an expensive medication to prevent HIV (Keller & Smith, 2011). This potentiality suggests that efforts to increase PrEP awareness and PrEP willingness must include discussions around access despite health insurance coverage, and also acknowledgement and efforts toward policy level changes in promoting health insurance access and affordability.

Discussions about, and sentiment toward Truvada as PrEP are complicated matters, affected by a variety of medical, behavioral, demographic, and social/structural factors (Mimiaga et al, 2009; Smith, Toledo, Smith, Adams, & Rothenberg, 2012; Perez-Figueroa, Kapadia, Barton, Eddy, & Halkitis, 2015). We therefore do not claim that this current form of analysis in and of itself is sufficient for understanding individuals’ conversations about, and sentiment toward Truvada; however, we do believe that the present method of using Twitter data to better understand these issues on the whole is low in cost and offers interesting insights to guide future intervention work. In short, we believe that this strategy is a key piece of working toward more effectively allocating resources and attention toward areas most in need to alleviate risk and disease.

Limitations

The present study’s findings should be interpreted in light of its limitations. First, the results are not generalizable to other areas of the U.S. not included in the MSAs for which AIDSVu had HIV prevalence rate and health insurance coverage data, nor are they generalizable to individuals who do not use Twitter or other social media sources to discuss or express their
attitudes toward Truvada as PrEP. Data were collected over a relatively short, six-month period, and do not reflect discussions or sentiment toward Truvada as PrEP outside that window of time; we did not investigate whether discussions or sentiment changed in any substantive way over the six-month period. Additional, longer studies should be performed to determine whether these factors do shift over time.

Conclusion

Results from this study suggest that it is feasible to use Twitter to identify Truvada as PrEP-related communications, determine probable locations for those communications, and then link qualities about those communications (i.e., frequency, sentiment) to area-level data for additional analyses, including HIV prevalence rates and health insurance coverage rates. This study calls for future research to understand other area-level factors that are associated with public discussions about, and sentiment toward, Truvada as PrEP. Additional research is also needed to determine the potential cost-effectiveness of using this approach to monitor and identify areas that are most in need of intervention. Social media data may serve as a way to bridge extant gaps between research and policy such that members of the two field may work together to identify and address areas that are most in need.
References


(PrEP) for HIV infection: results of a survey of HIV healthcare providers evaluating their knowledge, attitudes, and prescribing practices. *AIDS Patient Care and STDs, 27*(10), 553-559.


Table 1

Pearson Correlations (r) Between the Independent and Dependent Variables for the Geolocated Tweets in the U.S. MSAs of Interest

<table>
<thead>
<tr>
<th>Variable</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tweet frequency</td>
<td>0.391**</td>
<td>0.269**</td>
<td>0.07</td>
</tr>
<tr>
<td>2. Tweet sentiment</td>
<td>-</td>
<td>0.211*</td>
<td>0.178*</td>
</tr>
<tr>
<td>3. HIV prevalence rate</td>
<td>0.172*</td>
<td>-</td>
<td>-0.05</td>
</tr>
<tr>
<td>4. Health insurance coverage</td>
<td>0.223**</td>
<td>-0.05</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. * $p < 0.05$, ** $p < 0.01$. 
Table 2

**Fixed Effects Regressions Results of Factors Associated with Tweet Frequency and Tweet Sentiment**

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Outcome: Tweet Frequency</th>
<th></th>
<th></th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient (β)</td>
<td>Standard Error (SE)</td>
<td>p value</td>
<td></td>
</tr>
<tr>
<td><strong>HIV prevalence rate</strong></td>
<td>0.779</td>
<td>0.147</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td><strong>Health insurance coverage</strong></td>
<td>0.083</td>
<td>-0.140</td>
<td>0.571</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Outcome: Tweet Sentiment</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient (β)</td>
<td>Standard Error (SE)</td>
<td>p value</td>
<td></td>
</tr>
<tr>
<td><strong>HIV prevalence</strong></td>
<td>0.206</td>
<td>0.111</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td><strong>Health insurance coverage</strong></td>
<td>0.567</td>
<td>0.244</td>
<td>0.010</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 5: Discussion

The above-mentioned three studies, which together form the body of the present dissertation, must be interpreted not just separately, but also together.

First, each of these studies show that social and structural-level variables are key to establishing a more complete understanding of the HIV prevention and diagnosis landscape for this population. In each study, social and structural contextual variables emerged as statistically significant independent variables, while individual-level variables were either not significant, or became not significant after social and structural variables were added. These findings should not be interpreted as an attempt to undermine or minimize the importance of behavioral, demographic, and psychosocial characteristics, and interventions that address and consider these factors, but instead act as a call to develop more multi-level interventions which include variables at social/structural levels of influence, and to work more directly and more often with policymakers, rather than stopping at making generalized policy recommendations in scientific literature which are likely infrequently read and/or utilized by those outside of the field. The findings of these studies together ultimately support the theoretical frameworks that were used to guide the development of the present dissertation (Boerma & Weir, 2005), which posit that social/structural contexts are overwhelmingly strong promoters or hinderers of behavior, and without addressing these issues, behavioral change is likely to not sustain.

Second, these studies call upon us to establish a new definition of what it means to be at risk for HIV. Typically, entry criteria for studies involving issues surrounding HIV prevention in key populations (e.g., BMSM) center around recent sexual- or drug-related risk behaviors, and neglect to measure the extent to which neighborhood-level risk factors exist for these groups. This approach may be short-sighted, and ultimately will inhibit our ability to address risks at their root
causes. Two individuals who engage in similar sexual- and/or drug-related HIV risk behaviors in two neighborhoods of varying degrees of risk are not equivalently at risk for HIV. To accurately determine one’s HIV risk, neighborhood-level variables, such as HIV prevalence rate, access to services, neighborhood poverty, and urbanness/ruralness of the area, must be included.

Third, these studies show that there is a wealth of information available in data that has already been collected, and that these data can be deeply informative in guiding our decisions around where, when, and with whom we should intervene. Pulling from these sources is cost-effective, approachable, and allows scientists to perform multi-level analyses on datasets that were not originally intended for such a purpose, but from which we can gather rich social-ecological information about risk and disease acquisition.

While there was no evidence to support enacted stigma’s relation to HIV or STI cases, it was found to account for roughly 10% and 5% of HIV and STI cases in BMSM, respectively. This finding urges future research with larger sample sizes in this area, and calls for additional PAF calculations which consider looking at intersectional stigma to see whether multiple stigmas compound upon one another.

In Study 3, we found that HIV prevalence was positively associated with Truvada as PrEP conversation frequency, and positive attitudes toward Truvada as PrEP. However, in Study 1, HIV prevalence was found to be positively associated with testing positive for HIV. Therefore, it may be suggested that, while knowing about PrEP and feeling positively about PrEP is more common in areas where HIV prevalence is higher, knowing and feeling positively about PrEP does not mean that most individuals are taking, and being adherent to PrEP. While being aware of PrEP and feeling positively toward it is an important first step, these findings suggest that in areas with higher HIV prevalence, other barriers exist to PrEP uptake that must be addressed. Future research
warrants the exploration of PrEP uptake barriers beyond awareness and willingness to use, especially barriers at the social/structural levels.

In Study 3, we found that health insurance coverage was positively related to Truvada as PrEP sentiment, such that areas with higher insurance coverage had better sentiment towards Truvada, and areas with lower insurance coverage had worse sentiment towards Truvada. In Study 3, we found that neighborhood poverty was positively associated with testing positive for HIV, and in Study 1, we found that distance from services was positively associated with testing positive for HIV. Taken together, the findings of these studies suggest that, when access to care is low due to distance or health insurance coverage reasons, feelings toward HIV prevention tools like Truvada as PrEP are more negative, and also that acquiring HIV is more common. Additional work must be done in the future to better link individuals, especially key populations, to prevention and care services to promote HIV prevention behaviors like taking Truvada and reduce risk of HIV infection. Promoting these behaviors will likely require, to some degree, social/structural change around better access to care via transportation and more affordable health insurance coverage.