Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Katrece Outlaw Date

Who’s engaged? Using health information exchange to identify persons out of HIV care in Atlanta, GA

By

Katrece Outlaw
MPH

Epidemiology

_____

Travis Sanchez
Committee Chair

Who's engaged? Using health information exchange to identify persons out of HIV care in Atlanta, GA

By

Katrece Outlaw

B.A.
Washington University in St. Louis
2009

Thesis Committee Chair: Travis Sanchez, DVM, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2015

Abstract

Who's engaged? Using health information exchange to identify persons out of HIV care in Atlanta, GA

By Katrece Outlaw

New HIV transmission rates have remained stable in the United States for the last decade with an increasing burden of HIV diagnosis in the Southern states. A significant barrier to reducing HIV transmission rates is a failure to effectively engage persons living with HIV/AIDS (PLWHA) into HIV care. An emerging approach to address the lack of engagement in care is the development of HIV health information exchanges (HIE) between public health and health care providers. This study seeks to examine the utility of a HIE in identifying out of care patients who received non-HIV care at an Atlanta health center and to examine factors associated with the care engagement of the patients identified. Surveillance records from 8,350 PLWHA considered out of care by the Georgia Department of Public Health (GDPH) in 2014 were compared to 46,546 patient records from Saint Joseph's Mercy Care clinic (SJMC) from 2006 -2014. Patients were matched on the basis of name, date of birth and social security number. Bivariate and multivariable logistic regression analysis were used to examine demographic and clinical factors in this matched cohort associated with having never engaged in care or having at least one prior gap in care. Overall, 395 patients were identified as out of care who presented at SJMC. Of those, 109 (27.6%) had never engaged in care; of those who had engaged in care 219 (76.6%) had at least one gap in care. In multivariable analysis, patients who were homeless, aged 18-25 years or diagnosed for less than 10 years were more likely to have never engaged in care (PR= 1.63, 95% CI=1.21, 2.21; PR= 2.34, 95% CI=1.46, 3.77; PR= 3.17, 95% CI=1.95, 5.15). Bivariate analysis revealed that a diagnosis time of < 10 years and age 26-35 were associated with having a gap in care (PR= 0.87, 95% CI= 0.76-0.99; PR= 0.80, 95% CI= 0.65-0.98); however, no significant association was found in multivariable analysis (p=0.1647; p=0.0671). This study demonstrates that a HIE between the GDPH and SJMC could successfully identify patients out of care for HIV who present for other healthcare and provide opportunity for potential re-engagement.

Who's engaged? Using health information exchange to identify persons out of HIV care in Atlanta, GA

By

Katrece Outlaw

B.A.
Washington University in St. Louis
2009

Thesis Committee Chair: Travis Sanchez, DVM, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2015

Table of Contents

Chapter I: Background

HIV Epidemiology: United States, U.S. South and Georgia

HIV/AIDS care and prevention remain significant challenges for public health and the healthcare system in the United States. According to recent research from the Centers for Disease Control and Prevention, there are approximately 1.2 million persons living with HIV/AIDS (PLWHA) (1). The U.S. epidemic is a majority male epidemic as approximately 75% of PLWHA are male. The epidemic is also now one of middle age, with the largest age group of PLWHA between the ages of 40-49 (approximately 36%) (1). Although morbidity and mortality from HIV/AIDS has decreased with the advent of highly active antiretroviral therapy (HAART) in the late 1990's, the incidence of new HIV infections has remained stable at around 50,000 persons per year over the last decade (2). The epidemic does not affect the country uniformly but disproportionately affects certain regions and populations.

Statistics are particularly troubling in the Southern U.S. where rates of diagnosis and persons living with HIV/AIDS are higher than much of the country (3). The CDC defines the southern states as Alabama, Arkansas, Delaware, DC, Florida, Georgia, Kentucky, Louisianna, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia (3). The southern states accounted for nearly half of all new HIV diagnoses (48-49%) in the U.S. in 2011, while accounting for only 37% of the U.S. population (2, 4). In addition, 8 of the 10 states with the highest HIV diagnosis rates in 2010 were in the South.

Among the Southern states, the state of Georgia has some of the most troublesome statistics. As of 2012, there were approximately 57,000 PLWH in Georgia.

Demographic ratios for Georgians living with HIV/AIDS are similar to that of the nation as a whole with males accounting for approximately 75% of the HIV/AIDS population and the 45-54 year olds comprising the largest age group (1). Georgia had the 5th highest HIV diagnosis rate in the U.S. and the second highest diagnosis rate in the South in 2011 (4, 5).

HIV Care Continuum: United States, U.S. South and Georgia

Although advances in treatment with antiretroviral therapy (ART) have improved survival rates for people infected with HIV, rates of new infections in the U.S. have remained stable over the past decade. Current research indicates that a major barrier to reduction in HIV incidence is failure to effectively engage persons living with HIV in HIV care (6). Accordingly, the 2010 National HIV/AIDS Strategy includes a focus on efforts to improve linkage to and retention in care, with a 2013 Executive Order establishing a federal working group to focus efforts on improving outcomes along the HIV Care Continuum (7, 8).

The HIV Care Continuum is described as the sequential steps from diagnosis to optimal treatment (linkage, retention, prescription of ART and viral suppression). The definitions of the linkage and retention steps in the continuum are not always uniform, with no gold standard set for either of these terms (9, 10). Linkage is generally defined in most studies as one or more care visits within 3 months of HIV diagnosis (9, 11). Definitions of retention in care are much more varied (9, 10, 12). The Institute of Medicine and Health Resources and Services Administration (HRSA) HIV/AIDS Bureau define retention as 2 visits at least 90 days apart during a 12-month period (9, 13). The Department of Health and Human Services uses at least one visit during each 6-months of

a 24-month period with visits being at least 60 days apart. Other researchers have measured retention using the presence or absence of gaps, or a given time interval between visits, with intervals of 4-6 months defining a gap (9, 13). Other retention measures include missed visits and proportion of kept visits to scheduled visits (9). In many studies and for public health surveillance purposes, CD4 or viral load lab data have been used as a proxy for a clinic visit (14-16). Yehia et al found that the retention measure used by HRSA and the gap in care and visit constancy measures were significantly correlated and produce similar estimates of retention (10). However, strict attention to definitions and measures used is necessary when comparing estimates of retention across research studies.

The CDC produced estimates of the HIV Care Continuum for the years 2008 and 2011. The 2011 report, measuring engagement in care in 2008, found that of 1.2 million PLWHA in the U.S. 80% were diagnosed, 77% were linked to care, 51% were retained in care and only 35% were virally suppressed (14). A 2014 report, measuring engagement in care in 2011 found similar results. Of the PLWHA in 2011, 86% were diagnosed, only 40% were engaged in care, 37% were prescribed ART and only 30% were virally suppressed (15).

The Georgia Department of Public Health produced a report of the HIV Care Continuum in Georgia using the HRSA definition of retention in care. Of the PLWHA in Georgia in 2012, 82% were diagnosed, 44% were engaged in care, 28% had been prescribed ART and 32% were virally suppressed (16).

Disparities in HIV Transmission and Care

Rates of HIV prevalence and engagement along the care continuum are not uniform; significant disparities exist among racial minorities, particularly Black Americans, and also among young adults. Black Americans account for 41% of PLWHA and 44% of new HIV infections nationwide but account for only 12% of the U.S. population (2). Again the South represents an even greater disproportion among geographic regions. Black Americans account for 62% of all HIV diagnoses in the southern U.S. and account for 77% of all diagnoses in Georgia (2). In addition, Black Americans are consistently found to have the lowest rates of care along the HIV care continuum in both national and regional studies. In national studies, Blacks had lower linkage, retention in care, and viral suppression rates when compared to White Americans and other minority groups in the years 2009, 2010 and 2011 (13, 15, 17). Estimates were similar in the state of Georgia where Blacks had the lowest rates of retention and viral suppression among racial groups (16). Blacks also have higher rates of mortality for HIV/AIDS than any other racial group; accounting for 48% of HIV/AIDS related deaths in 2010 (2) and the highest age-adjusted HIV/AIDS death rate among racial groups.

HIV and lack of HIV care also disproportionately affect PLWHA ages 13-24 and 25-35. Nationally, those 35 years of age and younger accounted for 56% of all HIV infections in 2010 (2). PLWHA in the 35 years or younger age group were also less likely to be linked to care and have lower rates of viral suppression when compared with older age cohorts (15). Estimates are similar in the state of Georgia, where PLWHA aged 13-24 have the lowest rates of retention in care and viral suppression (16).

Factors associated with Engagement

Surveillance studies have identified race and age as potential factors associated with level of engagement in HIV care, as several studies have indicated that Black and Latino Americans and young persons (age 35 and younger) are less likely to be engaged at all levels of the care continuum than White Americans and older patients respectively (1, 14, 15). These findings are consistent with several cohort studies investigating factors associated with retention in care (17-20). Several other clinical and socio-contextual factors have been investigated as potential barriers to care including insurance status, mental health, substance abuse, CD4 and viral load counts. Studies by Giordano, Althoff, and Horstmann found that patients with a history of intravenous drug use were less likely to be retained in care (12, 20, 21). Hall and Horstmann et al also found that patients with higher CD4 counts were less likely to engage in care (19, 20). Overall, these findings suggest that public health and provider interventions will need to target a variety of populations in order to improve engagement in care.

Test and Treat

Unfortunately, the current state of engagement in care of PLWHA statistics does not support the success of the treatment as prevention strategy for prevention of HIV in the United States. The test-and-treat strategy has been touted by many researchers as a meaningful strategy to reduce HIV transmission in the United States (22, 23). The hypotheses guiding the test and treat strategy are 1) PLWHA who are tested and made aware of their HIV status are more likely to practice safer sex behaviors and 2) PLWHA who are engaged in care and receive early treatment with ART can achieve viral suppression and 3) PLWHA who have achieved viral suppression have lower levels of

virus in the blood and genital secretions and are less likely to transmit the virus sexually or via intravenous drug use (24-28).

There have been numerous studies in the last decade examining components of the test and treat strategy. Metsch et al demonstrated that PLWHA who were tested and linked to care had a reduction in sexual risk behaviors (24). Studies have shown that treatment with ART reduces HIV-1 RNA in genital secretions, including semen, cervical and vaginal secretions as well as ano-rectal mucosa (26, 29). Observational studies such as the work done by Baeten et al demonstrated a positive correlation between genital HIV-1 RNA and HIV transmission risk (26). Much of the research examining the relationship between this reduction in genital secretion of HIV and HIV transmission stems from observational studies of heterosexual, serodiscordant couples (29). To our knowledge only a single randomized control trial has reported results on this relationship to date (30). This single RCT conducted by the NIH found a 96% reduction in transmission with ART use in combination with condoms and risk reduction counseling. Overall, the body of current research demonstrates a consensus that ART is associated with decreased risk of transmission in serodiscordant couples (30, 31).

Mathematical modeling has also been used to demonstrate the utility of the test and treat model to prevent HIV transmission (23, 30). Many of these models conclude that test and treat is a viable option to substantially reduce HIV transmission; however, the success of the strategy depends on both widespread HIV testing and ART use. A recent study by Skarbinski et al examined transmission rates at each point in the HIV care continuum and found that persons undiagnosed with HIV or diagnosed but not retained in care accounted for 91.5% of all transmissions in the U.S. in 2009 (32). Thus efforts to

improve diagnosis and subsequent engagement in care are necessary to aid prevention efforts and decrease transmission. It is imperative that programs and initiatives are in place to link diagnosed individuals to care and to keep those individuals engaged in care (15).

Health Information Exchange and HIV Care

An emerging approach to address the lack of engagement in HIV care is the use of electronic health information exchange (HIE) between health care delivery systems. HIE has been described as "the process of electronic multi-directional transfer of identifiable, patient-level information between different organizations" (33) and has been used in primary care, emergency department and inpatient settings in attempts to improve coordination of care between clinical providers and other members of the health care system, to improve patient care efficiency and to improve patient outcomes (34, 35). In the context of HIV care, studies of HIE in Los Angeles and North Carolina have been shown to improve clinician access to patient care information, patient quality of care, maintenance of patient care guidelines, and cost-effectiveness (34, 36, 37). Bi-directional exchange of laboratory data between lab facilities and electronic health record (EHR) improved timeliness in adjusting ART regimen (38). Few studies have investigated HIE use in HIV care and its effects on patient retention and outcomes.

In 2007, HRSA developed an initiative to investigate the utility of HIE in improving engagement in HIV care (33, 39). The initiative provided support for six sites to implement HIE interventions of various structures for PLWHA in underserved communities. The results of this initiative demonstrate promising results for improvement of retention in HIV care. Overall, the study found that 91% of patients

were retained in care (retention defined as at least one primary care visit or CD4/viral load lab within the 6 month study period), 79% were receiving antiretroviral therapy and 52% had reached undetectable viral load.

Of particular interest, two sites within the study implemented HIEs using public health surveillance data to improve engagement. The HIE implemented in Louisiana sought to identify PLWH who are considered out-of-care (OOC) who were seen within the public Louisiana healthcare system for other health reasons in an effort to re-engage that person for HIV services (33, 39). The Louisiana HIE (LAPHIE) successfully identified > 400 OOC patients with 62% of whom had at least 1 HIV specialty care visit at 1-year follow-up (39, 40). The study demonstrated a clear utility of HIE in alerting providers of OOC patients and presents a model that can be replicated in other regions and health systems. In light of the recent law permitting the release of confidential HIV surveillance information to health care providers (41), the Georgia Department of Public Health (GDPH) seeks to implement a similar HIE in order to improve engagement in HIV care in Georgia. Through this study we seek to demonstrate the utility of a HIV HIE within Atlanta, GA and also identify factors associated with engagement, which may inform efforts to re-engage patients identified through the exchange. To our knowledge no study to date has investigated factors associated with engagement in care among those identified via a HIE. Georgia is a state with a significant HIV burden and low percentages of engagement on the HIV care continuum. This study would help to inform Georgia public health officials and clinical providers on opportunities to re-engage these individuals for proper linkage to and retention in care.

Chapter II: Manuscript

Who's engaged? Using health information exchange to identify persons out of HIV care in Atlanta, GA

Katrece Outlaw, Travis Sanchez, A. Eugene Pennisi

Abstract:

New HIV transmission rates have remained stable in the United States for the last decade with an increasing burden of HIV diagnosis in the Southern states. A significant barrier to reducing HIV transmission rates is a failure to effectively engage persons living with HIV/AIDS (PLWHA) into HIV care. An emerging approach to address the lack of engagement in care is the development of HIV health information exchanges (HIE) between public health and health care providers. This study seeks to examine the utility of a HIE in identifying out of care patients who received non-HIV care at an Atlanta health center and to examine factors associated with the care engagement of the patients identified. Surveillance records from 8,350 PLWHA considered out of care by the Georgia Department of Public Health (GDPH) in 2014 were compared to 46,546 patient records from Saint Joseph's Mercy Care clinic (SJMC) from 2006 -2014. Patients were matched on the basis of name, date of birth and social security number. Bivariate and multivariable logistic regression analysis were used to examine demographic and clinical factors in this matched cohort associated with having never engaged in care or having at least one prior gap in care. Overall, 395 patients were identified as out of care who presented at SJMC. Of those, 109 (27.6%) had never engaged in care; of those who had engaged in care 219 (76.6%) had at least one gap in care. In multivariable analysis, patients who were homeless, aged 18-25 years or diagnosed for less than 10 years were more likely to have never engaged in care (PR= 1.63, 95% CI=1.21, 2.21; PR= 2.34, 95% CI=1.46, 3.77; PR= 3.17, 95% CI=1.95, 5.15). Bivariate analysis revealed that a diagnosis time of < 10 years and age 26-35 were associated with having a gap in care (PR= 0.87, 95% CI= 0.76-0.99; PR= 0.80, 95% CI= 0.65-0.98); however, no significant association was found in multivariable analysis (p=0.1647; p=0.0671). This study demonstrates that a HIE between the GDPH and SJMC could successfully identify patients out of care for HIV who present for other healthcare and provide opportunity for potential re-engagement.

INTRODUCTION

Approximately 1.2 million people in the United States are living with HIV/AIDS. The epidemic in the U.S. continues to disproportionately affect specific regions and populations within the country as the southern U.S. and racial minority groups have higher rates of HIV/AIDS diagnoses than their northern and white counterparts respectively. HIV is particularly prevalent in Georgia with the 5th highest HIV diagnosis rate in the country and the 2nd highest diagnosis rate in the South (4, 5). A substantial amount of research now focuses effective interventions to reduce HIV transmissions.

Advances in treatment with HAART have improved survival rates for people infected with HIV; however, efforts in HIV prevention have not been as successful as rates of new infections in the U.S. have remained stable over the past decade (2). Failure to effectively link and engage PLWHA in HIV care represents a significant challenge for prevention efforts to reduce HIV incidence (6). PLWHA who are engaged in care and receiving ART have lower viral loads and reduced risk of HIV transmission (22, 30). However, most PLWHA do not reach viral suppression according to studies describing the HIV Care Continuum, the sequential steps from diagnosis to optimal treatment. Recent studies reveal that <40% of PLWHA are engaged in care and receiving ART and only 30% are virally suppressed (15). In addition young persons (<35 years of age) and African-Americans are less likely to be linked to, or retained in care (13, 15). Accordingly the National HIV/AIDS Strategy includes a focus on efforts to accelerate progress along the HIV Care Continuum (11).

The "test-and-treat" model, held by many researchers as an effective strategy for HIV prevention, is in many ways dependent upon improvement in the HIV Care Continuum. The test and treat model promotes universal HIV testing, coupled with early

initiation of HAART for those diagnosed (23, 28). This strategy is guided by research that has found that diagnosed PLWHA have less risky sexual behaviors and that those treated with HAART have reduced the amount of HIV virus in blood plasma and genital secretions and thus reduced risk of transmitting the virus (24, 26, 30, 31). A study investigating transmission rates at each step in the care continuum found that PLWHA who were undiagnosed or diagnosed but not retained in care accounted for 91.5% of all transmissions in 2009 (32). Thus it is apparent that to implement a strategy of treatment as prevention for HIV, efforts must be made to improve linkage to and retention in care.

The development of HIV health information exchanges (HIE) between public health and health care delivery systems is a promising strategy to address the lack of engagement in care. In 2007, HRSA funded an initiative to investigate the utility of HIEs in improving linkage to HIV care. Included in that study is an HIE implemented in Louisiana which sought to identify PLWHA who are considered out-of-care (OOC) who were seen within the public Louisiana healthcare system for other health reasons in an effort to re-engage that person for HIV services (33). The Louisiana HIE, identified > 300 OOC patients, >60% of whom were effectively re-engaged in HIV care at 1 year follow-up (39). In light of the recent law permitting the release of confidential HIV surveillance information to health care providers, the Georgia Department of Public Health (GDPH) seeks to implement a similar HIE in order to improve engagement in HIV care in Georgia (41).

The purpose of this study is to investigate the utility of a HIV HIE within Atlanta, GA and also to identify factors associated with engagement. We hypothesize that a substantial number of PLWHA in Georgia who are out of care can be identified via HIE

between the GDPH and a local community clinic. We further hypothesize that Blacks and younger PLWHA will be more likely to lack engagement in care. To our knowledge no study to date has investigated factors associated with engagement in care among patients identified via a HIE. The results of this study could help to inform GDPH officials on opportunities to re-engage these individuals for improved linkage to and retention in care.

METHODS

Objectives

- To examine how many PLWH who are considered out-of-care (OOC) for HIV could have been identified via a health information exchange between GDPH and Saint Joseph's Mercy Care (SJMC) in Atlanta, Georgia based on GDPH Enhanced HIV/AIDS Reporting System (eHARS) surveillance data and SJMC electronic medical record (EMR) data.
- To determine factors associated with patients never engaging in care as compared to patients who have gaps in care (cyclical engagement) among those persons identified.

Hypotheses

- Approximately 2% of individuals OOC for HIV will be identified via an HIE between the GDPH and SJMC
- Patients of Black or Hispanic race as well as young patients will be more likely to have never engaged in care or have gaps in care.

Study Population

We analyzed patient data collected by the GDPH for routine HIV/AIDS surveillance as well as patient data from St. Joseph's Mercy Care (SJMC), a federally qualified health center and patient centered medical home in Atlanta, GA. There were 84,800 SJMC clinic patient records and 89,526 GDPH Enhanced HIV/AIDS Reporting System (eHARS) patient records reviewed for this study. SJMC and eHARS records prior to 2006 were not included in the study due to inconsistent HIV laboratory reporting prior to this time period. SJMC records from 2006 to 2014 were included for maximal potential for identifying matched patients. April 21, 2014 was the last date of clinic visit data provided by SJMC.

Patients from the eHARS database were included in match identification analysis if the patient was alive as of April 21, 2014, had a Georgia address, had a first and last name, had at least one independent lab document confirming HIV diagnosis and was considered out of care (OOC) as of April 21, 2014. For the purposes of this study, OOC is defined as a lack of CD4+ T-lymphocyte count (CD4) or HIV viral load (VL) lab data in eHARS for a period of 455 days or more preceding the current date. This means that the person had either never had a recorded lab date or had not had a lab test since January 8, 2013. Patients are generally considered out of care by the GDPH if missing laboratory data for the previous one year. The period of 455 days was chosen for this study to allow for an additional 90 days of possible reporting delay.

Patients identified from the SJMC patient record and eHARS were included in the sample for analysis if matched on the basis of name, sex, and date of birth as recorded in HIV surveillance eHARS database and SJMC electronic medical record (EMR). The last

4 digits of the social security number was also used to identify matched individuals if recorded in the eHARS database.

Direct patient identifiers were used only for the purposes of matching OOC patients to SJMC clinic records. Once matched, data was de-identified for further analysis. This study involved the use of protected health information (PHI), some of which may be sensitive should inadvertent release occur. The data set containing PHI was created by GDPH staff whom have access to this information through their regular roles in the provision and monitoring of the Georgia HIV surveillance program and Health Information Exchange. This data was accessed only on secured servers at secured offices at the GDPH. This study did not involve informed consent as this research involves no greater than minimal risk to human subjects, it involves secondary use of data collected for public health surveillance, and no information would be derived from this evaluation that would directly benefit the subjects. The study was approved by the Emory University Institutional Review Board.

Outcome and Variable Definitions

Outcomes of interest in this analysis were 1) never in care (NIC) 2) out of care with prior gaps in care and 3) out of care with no prior gaps in care. This analysis examined the differences between those never in care and previously engaged in care (with prior gap and without). We also examined the difference between those previously engaged in care with prior gaps versus those previously in care without gaps.

For the purposes of this study patients were defined as never engaged in care if the most recent CD4 or VL date was within three months of the date of diagnosis. Patients with no lab document data recorded in eHARS were also considered never

engaged in care. Patients were defined as having a gap in care if time between any two lab dates exceeded 180 days. Similar definitions for retention/gap have been used in previous research (9, 13, 42).

Eight demographic and clinical variables were examined as factors contributing to the outcomes: Race, current age, gender, risk, time since HIV diagnosis, CD4 count (first and last) and housing status. Race was categorized into Black/African American, White, Hispanic and Other. Current age was categorized as 18-25, 26- 35, and > 35 years. Gender was categorized as male, female, or transgender. All transgender patients were male to female transition. Risk or transmission category at diagnosis was dichotomized into men who have sex with men (MSM) and other. MSM included those with only MSM as a risk factor and those with both MSM and intravenous drug use (IDU) as risk factors. Time since HIV diagnosis was also dichotomized into greater than or less than 10 years. Both most recent and first CD4 counts were used in analyses. First CD4 count reflected the CD4 lab value measured soonest after HIV diagnosis and was considered a possible factor for never engaged in care. Most recent CD4 count reflected the CD4 lab value recorded at the most recent lab visit and was considered a possible factor for having a gap in care. Both CD4 variables were dichotomized to reflect whether the count was > or $\leq$ 200 cells/mm^3. Housing status at diagnosis was categorized into either homeless or not homeless at time of last SJMC clinic visit. Homeless included persons who reported living with others, in a shelter/boarding house, or other structured facility (substance abuse facility, hospital, or jail). Date data with missing values for month or day were assigned January for the month and the 15^{th} for the day respectively for the purposes of

variable creation and calculation during analysis. For example, if the date recorded was 05/2008, the date used in the analysis was 05/15/2008.

Data Analyses

Basic descriptive analyses were used to estimate the proportion of socio-demographic and clinical variables of the overall matched samples as well as the three outcome groups. First CD4 count was found to be missing for the majority of the sample with the NIC outcome and was excluded from further analysis as missing recorded CD4 or VL lab data was used to define this outcome. Bivariate and multivariable logistic regression analyses were used to examine factors associated with the engagement outcomes (never engaged in care or having a gap in care). Variables found to be significant at a level of 0.05 in the bivariate analyses were tested for inclusion in the final multivariable logistic regression models. Prevalence ratios were calculated using SAS callable SUDAAN logistic regression with predictive margins. Factor variables were examined for possible correlation using Pearson correlation analyses for comparison of continuous variables and chi-square analyses for dichotomous and categorical variables. Factor variables were also examined for possible interaction using stratified logistic regression as well as examination of the significance of two-way interaction terms within the models. All statistical analyses were performed using SAS 9.4 (Cary, N.C.).

RESULTS

A total of 89,526 HIV-positive patients were recorded in the GDPH eHARS system, including all reported diagnoses from reporting counties in the state of Georgia. 81,176 patients were removed from analysis, as they did not meet selection criteria for inclusion in the GDPH out of care list. A total of 84,833 patient records were obtained

from SJMC for patient visits between 2006 and 2014. Of those, 38,254 were removed due to missing encounter date. An additional 33 records were removed due to missing last name, first name or date of birth values.

A total of 395 HIV-positive patients were identified as currently out of care and who presented at SJMC for health care between 2006 and 2014. Table 1 shows the socio-demographic and clinical data for the sample. The majority of the sample were male (70.9%) and Black/African-American (79.2%). Six patients (1.5%) in the sample were male-to-female transgender persons. The median age for the sample was 44 years with 72.9% of patients in the sample greater than age 35. The majority of patients had MSM as the risk factor for HIV infection. Forty-three percent of patients had their HIV diagnosis more than 10 years prior and the median of first CD4 count was 286 cells/mm^3. Correlation analysis of factor variables revealed that older age was associated with homelessness (X^2=8.35, p=0.0154). Older age was also strongly associated with longer time since diagnosis (X^2=52.03 p <0.0001). Shorter time since diagnosis was also highly associated with having a missing recent CD4 count (X^2=45.94, p <0.0001). Analysis did not demonstrate any other significant correlation between variables.

Never Engaged in Care

Overall, 109 identified OOC patients (27.9%) were never engaged in care (Table 1). Fifty-eight NIC patients (53.2%) were missing values for first CD4 count variable and this variable excluded from further analysis (Table 1). In a comparison of those NIC versus those receiving some prior HIV care, never having been engaged in care was more prevalent among patients whom were homeless, less than 35 years of age, and who were diagnosed for 10 years or less (Table 2). The prevalence of NIC was 2 greater among

homeless patients when compared to housed patients (PR=1.63, 95% CI=1.21, 2.21). The prevalence of NIC was also 2 times greater among patients age 18-25 when compared to patients over the age of 35 (PR=2.34, 95% CI=1.46, 3.77) and 3 times greater among patients diagnosed for 10 years or less when compared to those diagnosed >10 years (PR= 3.17, 95% CI=1.95-5.15). Race, gender and risk category were not significantly associated with never having been engaged in care. Age was not a significant factor for patients who were homeless but was strongly significant for those who were housed. When compared to housed patients ≥35 years of age, the prevalence of NIC was 5 times greater among housed patients age 18-25 and 2 times greater among housed patients age 26-34 (PR=5.10, 95% C =3.17, 8.20; PR=2.00, 95% CI=1.18, 3.39). The interaction term, however, was not significant in the final model (p=0.2129).

Gap in Care

Of those who had established care, 219 patients (76.6%) had at least one prior gap in care. Of the patients who had at least one gap in care, 88 (40%) had only 1 gap, 67 (30.5%) had 2 gaps in care and 65 (29.5%) had greater than 2 gaps in care. In the bivariate analysis, only a diagnosis time greater than 10 years and current age were associated with having a gap in care (Table 3). Neither factor was significant in the multivarable model.

DISCUSSION

Of 8,350 patients considered OOC in the state of Georgia as of 2014, 395 were identified as having patient encounters at St. Joseph's Mercy Care clinic between 2006 and 2014. The majority of OOC patients identified via this study had been linked to care at some time period during their diagnosis but had at least one gap in care. Of 395 OOC patients, 109 had never engaged in care. Among patients who had never engaged in care,

housing status, current age, and first CD4 count emerged as factors associated with lack of engagement in multivariate analysis.

Persons who were homeless at time of diagnosis demonstrated more difficulty engaging in HIV care. Homeless patients may have increased barriers to accessing care due to housing instability. Patients may be harder to contact for linkage due to lack of reliable address and may also prioritize the need to secure housing over clinic appointments. No studies reviewed for this research particularly addressed homelessness as a factor for engagement in care. However, research has found that other factors which may correlate with homelessness, such as illicit drug use and mental illness, are associated with delayed entry into care (21, 43, 44). Providers should be prepared to help out of care patients address structural barriers such as homelessness via case management or referral to other resources in order to help those patients to engage and be retained in care.

Patients 18-25 years of age were also more likely to never engage in HIV care. These findings are consistent with prior cohort studies and surveillance data, which found younger age to be associated with higher rates of delay or lack of engagement in care (15, 17, 19). These results suggest that providers should pay particular attention to young patients identified by HIE as out of care. These patients may have less contact with healthcare in general and have a greater sense of invulnerability than older patients and may require additional counseling and education to engage them in care (17, 21).

Patients with a shorter duration of diagnosis were also less likely to enter HIV care. These findings may reflect a psychosocial barrier in care whereby patients do not engage in care until they fall ill. Those who have been diagnosed for less time may have

less opportunity to develop an illness that would cause them to seek care. There may also be an unmeasured confounder such as testing reason or testing facility that influenced the results of this study. Patients who are tested for the first time or for reasons other than perceived high risk are more likely to delay entry into care (44, 45). Patients diagnosed in the last decade may have been more likely to have been tested under these circumstances due to the revised CDC recommendations for increased testing in 2006 (46). These findings differ from studies by Giordano and Marks et al, which did not reveal any significant association between time since diagnosis and entry into care (21, 47).

Notably, our study did not identify race or gender as significant factors associated with engagement in care, as these demographic factors were strongly associated with engagement in previous studies (17, 19, 21, 48). In particular, regional and national surveillance reports and cohort studies have identified Black Americans as significantly more likely to have delayed care, not entered into care, or not maintained care (2, 15, 17, 19, 21, 48). However, our study contained a much more homogeneous demographic sample than the majority of these previous studies with nearly 80% of the population identifying as Black/African American and only approximately 10% identifying as White. In this more homogeneous patient population, race could be more difficult to distinguish as an engagement factor. Other factors unmeasured in this study such as socioeconomic and insurance status, stigma and structural barriers could be investigated as potential contributors to lack of engagement in this population.

Studies by Konkle et al and Sprague et al examined some of these factors via structured interview (49, 50). Patients in these studies reported feeling well, privacy concerns, and lack of transportation as major reasons for delaying their entry into care.

Providers should educate patients on disease course and address concerns of confidentiality when attempting to re-engage persons identified as out of care. Previous studies also found more recent diagnosis to be associated with improved rates of retention (17, 19, 20). These findings are consistent with our study findings of a strong correlation between diagnosis time and presence of a gap in care. Patients with a greater amount of time since initial diagnosis have more opportunity to fall out of care and develop gaps between visits. Our study failed to detect any additional factors that may be associated with having gaps in care. Previous studies have found age, sex, HIV risk category, and CD4 count to be associated with retention in care with older age, MSM, and lower CD4 counts being associated with higher likelihood of engagement in care (12, 17, 19, 20). These studies used the DHHS definition for retention, which has been found to correlate with the gap measure (10). This study's lack of findings could reflect a lack of ability to detect significant variance due to small sample size. In addition, there may be factors that remained unmeasured in this study relating to the development of gaps in care that could further characterize patients with this pattern of engagement.

There are several limitations to this study. This study utilizes SJMC patient encounter data from patients seen from 2006 – 2014. Utilization of this date range allows us to estimate the maximum number of persons currently out of care who could possibly be identified via clinic visit at SJMC. However, this estimation relies on the assumption that patients consistently receive healthcare at SJMC and will return to the clinic for future care in the event that their last encounter was several years past. In addition there may be unmeasured factors attributable to the care at SJMC that are associated with patients' out of care status that cannot be examined. This study is also limited by the use

of surveillance data to determine factor and outcome variables. Many socioeconomic and structural factors that may serve as barriers to care are not recorded or measured in surveillance data. Outcome measures of engagement and retention are also limited by use of surveillance data as CD4 or viral load lab records are used as a proxy for HIV care visits. This can result in overestimation of lack of engagement if patients' labs are not reported or the patient moves out of the reporting area (44).

In conclusion, this study is the first to attempt to identify patients out of care for HIV via a HIE between the GDPH and a local community clinic. This study is also the first to further examine the factors associated with lack of engagement in those patients identified. This study demonstrates that approximately 5% of out of care PLWHA can be identified via such a HIE, particularly if the patient population receives consistent care at SJMC. Identification of these patients at the time of the patient encounter can give providers an invaluable opportunity to reengage those individuals into care. Providers should pay particular care to counsel and address the concerns of younger patients, those who are homeless, and those who have been diagnosed for 10 years or less. This study found these patients to be more likely to have never engaged in care and previous studies demonstrate that these patients are also less likely to remain in care once engaged. Successful development and implementation of a HIE has the potential not only to re-engage out of care patients but also reduce disparities in HIV care by targeting vulnerable populations.

REFERENCES

1. Center for Disease Control and Prevention. HIV Surveillance Report. 2013.
2. Kaiser Family Foundation. The HIV/AIDS Epidemic in the United States. 2014. (http://kff.org/hivaids/fact-sheet/the-hivaids-epidemic-in-the-united-states/ - footnote-107662-2). (Accessed 2014).
3. Center for Disease Control and Prevention. HIV and AIDS in the United States by Geographic Distribution. 2012.
4. Reif SS, Whetten K, Wilson ER, et al. HIV/AIDS in the Southern USA: a disproportionate epidemic. *AIDS care* 2014;26(3):351-9.
5. Center for Disease Control and Prevention. 2013 - Georgia State Health Profile. 2013.
6. Gardner EM, McLees MP, Steiner JF, et al. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2011;52(6):793-800.
7. The White House. National HIV/AIDS Strategy for the United States. 2010.
8. The White House. Executive Order - HIV Care Continuum Initiative. 2013.
9. Mugavero MJ, Amico KR, Horn T, et al. The state of engagement in HIV care in the United States: from cascade to continuum to control. *Clinical infectious diseases* 2013;57(8):1164-71.
10. Yehia BR, Fleishman JA, Metlay JP, et al. Comparing different measures of retention in outpatient HIV care. *AIDS* 2012;26(9):1131-9.

11. The White House. National HIV/AIDS Strategy – Improving Outcomes: Accelerating Progress Along the HIV Care Continuum. 2013.
12. Althoff KN, Rebeiro P, Brooks JT, et al. Disparities in the quality of HIV care when using US Department of Health and Human Services indicators. *Clinical infectious diseases* 2014;58(8):1185-9.
13. Mugavero MJ, Davila JA, Nevin CR, et al. From Access to Engagement: Measuring Retention in Outpatient HIV Clinical Care. *AIDS patient care and STDs* 2010;24(10):607-13.
14. Vital signs: HIV prevention through care and treatment--United States. *MMWR Morbidity and mortality weekly report* 2011;60(47):1618-23.
15. Bradley H, Hall HI, Wolitski RJ, et al. Vital Signs: HIV diagnosis, care, and treatment among persons living with HIV--United States, 2011. *MMWR Morbidity and mortality weekly report* 2014;63(47):1113-7.
16. Georgia Department of Public Health. HIV/AIDS Epidemiology Program HIV Care Continuum Surveil¬lance Report, Georgia, 2012. 2012.
17. Fleishman JA, Yehia BR, Moore RD, et al. Establishment, retention, and loss to follow-up in outpatient HIV care. *Journal of acquired immune deficiency syndromes (1999)* 2012;60(3):249-59.
18. Tripathi A, Gardner LI, Ogbuanu I, et al. Predictors of time to enter medical care after a new HIV diagnosis: a statewide population-based study. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv* 2011;23(11):1366-73.

19. Hall HI, Gray KM, Tang T, et al. Retention in Care of Adults and Adolescents Living With HIV in 13 US Areas. *Jaids-Journal of Acquired Immune Deficiency Syndromes* 2012;60(1):77-82.

20. Horstmann E, Brown J, Islam F, et al. Retaining HIV-infected patients in care: Where are we? Where do we go from here? *Clinical infectious diseases* 2010;50(5):752-61.

21. Giordano TP, Visnegarwala F, White AC, Jr., et al. Patients referred to an urban HIV clinic frequently fail to establish care: factors predicting failure. *AIDS care* 2005;17(6):773-83.

22. Dodd PJ, Garnett GP, Hallett TB. Examining the promise of HIV elimination by 'test and treat' in hyperendemic settings. *AIDS* 2010;24(5):729-35.

23. Granich RM, Gilks CF, Dye C, et al. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet* 2009;373(9657):48-57.

24. Metsch LR, Pereyra M, Messinger S, et al. HIV transmission risk behaviors among HIV-infected persons who are successfully linked to care. *Clinical infectious diseases* 2008;47(4):577-84.

25. Kalichman SC, Cherry C, Amaral CM, et al. Adherence to antiretroviral therapy and HIV transmission risks: implications for test-and-treat approaches to HIV prevention. *AIDS patient care and STDs* 2010;24(5):271-7.

26. Baeten JM, Kahle E, Lingappa JR, et al. Genital HIV-1 RNA predicts risk of heterosexual HIV-1 transmission. *Science translational medicine* 2011;3(77):77ra29.

27. Donnell D, Baeten JM, Kiarie J, et al. Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. *Lancet* 2010;375(9731):2092-8.

28. Mugavero MJ, Amico KR, Westfall AO, et al. Early retention in HIV care and viral load suppression: implications for a test and treat approach to HIV prevention. *Journal of acquired immune deficiency syndromes (1999)* 2012;59(1):86-93.

29. Cohen MS, Gay CL. Treatment to prevent transmission of HIV-1. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2010;50 Suppl 3:S85-95.

30. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *The New England journal of medicine* 2011;365(6):493-505.

31. Anglemyer A, Rutherford GW, Horvath T, et al. Antiretroviral therapy for prevention of HIV transmission in HIV-discordant couples. *The Cochrane database of systematic reviews* 2013;4:CD009153.

32. Skarbinski J, Rosenberg E, Paz-Bailey G, et al. Human Immunodeficiency Virus Transmission at Each Step of the Care Continuum in the United States. *JAMA internal medicine* 2015.

33. Shade SB, Chakravarty D, Koester KA, et al. Health information exchange interventions can enhance quality and continuity of HIV care. *International journal of medical informatics* 2012;81(10):e1-9.

34. Proeschold-Bell RJ, Belden CM, Parnell H, et al. A randomized controlled trial of health information exchange between human immunodeficiency virus institutions. *Journal of public health management and practice : JPHMP* 2010;16(6):521-8.

35. Hincapie A, Warholak T. The Impact of Health Information Exchange on Health Outcomes. *Applied Clinical Informatics* 2011;2(4):499-507.

36. Messer LC, Parnell H, Huffaker R, et al. The development of a health information exchange to enhance care and improve patient outcomes among HIV+ individuals in rural North Carolina. *International journal of medical informatics* 2012;81(10):e46-55.

37. Myers JJ, Koester KA, Chakravarty D, et al. Perceptions regarding the ease of use and usefulness of health information exchange systems among medical providers, case managers and non-clinical staff members working in HIV care and community settings. *International journal of medical informatics* 2012;81(10):e21-9.

38. Bell DS, Cima L, Seiden DS, et al. Effects of laboratory data exchange in the care of patients with HIV. *International journal of medical informatics* 2012;81(10):e74-82.

39. Herwehe J, Wilbright W, Abrams A, et al. Implementation of an innovative, integrated electronic medical record (EMR) and public health information exchange for HIV/AIDS. *Journal of the American Medical Informatics Association : JAMIA* 2012;19(3):448-52.

40. Magnus M, Herwehe J, Gruber D, et al. Improved HIV-related outcomes associated with implementation of a novel public health information exchange. *International journal of medical informatics* 2012;81(10):e30-8.

41. Burke D, Unterman R. Georgia General Assembly. Senate Bill 342. *Georgia General Assembly* (http://www.legis.ga.gov/legislation/en-US/display/20132014/SB/342) (Accessed 27 Aug. 2014).

42. Gardner EM, Daniloff E, Thrun MW, et al. Initial linkage and subsequent retention in HIV care for a newly diagnosed HIV-infected cohort in Denver, Colorado. *Journal of the International Association of Providers of AIDS Care* 2013;12(6):384-90.

43. Toth M, Messer LC, Quinlivan EB. Barriers to HIV Care for Women of Color Living in the Southeastern US Are Associated with Physical Symptoms, Social Environment, and Self-Determination. *AIDS patient care and STDs* 2013;27(11):613-20.

44. Jenness SM, Myers JE, Neaigus A, et al. Delayed entry into HIV medical care after HIV diagnosis: Risk factors and research methods. *Aids Care-Psychological and Socio-Medical Aspects of Aids/Hiv* 2012;24(10):1240-8.

45. Reed JB, Hanson D, McNaghten AD, et al. HIV Testing Factors Associated with Delayed Entry into HIV Medical Care among HIV-Infected Persons from Eighteen States, United States, 2000-2004. *AIDS patient care and STDs* 2009;23(9):765-73.

46. Branson BM, Handsfield HH, Lampe MA, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings.

MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports / Centers for Disease Control 2006;55(RR-14):1-17; quiz CE1-4.

47. Marks G, Gardner LI, Craw J, et al. Entry and retention in medical care among HIV-diagnosed persons: a meta-analysis. *AIDS* 2010;24(17):2665-78.

48. Mugavero MJ, Lin HY, Allison JJ, et al. Failure to establish HIV care: Characterizing the "no show" phenomenon. *Clinical Infectious Diseases* 2007;45(1):127-30.

49. Konkle-Parker DJ, Amico KR, Henderson HM. Barriers and facilitators to engagement in HIV clinical care in the Deep South: results from semi-structured patient interviews. *The Journal of the Association of Nurses in AIDS Care : JANAC* 2011;22(2):90-9.

50. Sprague C, Simon SE. Understanding HIV care delays in the US South and the role of the social-level in HIV care engagement/retention: a qualitative study. *International journal for equity in health* 2014;13:28.

51. Higa DH, Marks G, Crepaz N, et al. Interventions to improve retention in HIV primary care: a systematic review of U.S. studies. *Current HIV/AIDS reports* 2012;9(4):313-25.

52. Ulett KB, Willig JH, Lin HY, et al. The therapeutic implications of timely linkage and early retention in HIV care. *AIDS patient care and STDs* 2009;23(1):41-9.

Tables

Table 1. Demographic and Clinical Characteristics for Individuals Out of Care for HIV Services Receiving Medical Services at St Joseph's Mercy Care From 2006 - 2014, N=395

Variable	Total	Never in Care*	Gap in Care**	No Gap in Care
Overall	395	109 (27.6)	219 (55.4)	67 (17)
Race; N(%)				
Black	313 (79.2)	86 (78.9)	176 (80.4)	51 (76.1)
White	39 (9.9)	11. (10.1)	18 (8.2)	10 (14.9)
Hispanic	25 (6.3)	7 (6.4)	16 (7.3)	2 (.04)
Other	17 (4.3)	4 (3.7)	9 (4.1)	4 (.06)
Missing	1 (0.3)	1 (0.9)	0	0
Gender; N(%)				
Male	280 (70.9)	74 (67.9)	157 (71.7)	49 (73.1)
Female	55 (13.9)	10 (9.2)	35 (16.0)	10 (14.9)
Transgender	6 (1.5)	0	6 (2.7)	0
Missing	54 (13.7)	25 (22.9)	21 (9.6)	8 (11.9)
Housing; N(%)				
Not Homeless	235 (59.5)	53 (48.6)	139 (63.5)	43 (64.2)
Homeless	160 (40.5)	56 (51.4)	80 (36.5)	24 (35.8)
Age; N(%)				
18-25	17 (4.3)	12 (11.0)	4 (1.8)	1 (0.02)
26-35	90 (22.8)	32 (29.4)	37 (16.9)	21 (31.3)
>35	288 (72.9)	65 (59.6)	178 (81.3)	45 (67.2)
Age; Median (SD)[§]	44.0 (11.0)	40.0 (11.3)	47.0 (10.4)	43.0 (10.7)
Risk; N(%)				
MSM^	273 (69.1)	72 (66.1)	152 (69.4)	49 (73.1)
Other	122 (30.9)	37 (33.9)	67 (30.6)	18 (26.9)
Diagnosis Time° N(%)				
> 10 yrs	170 (43.0)	18 (16.5)	124 (56.6)	28 (41.8)
≤ 10 yrs	225 (57.0)	91 (83.5)	95 (43.4)	39 (58.2)
1st CD4[£]; Median (SD)	286.0 (274.4)	332.0 (292.2)	272.0 (268.3)	290.0 (273.5)
1st CD4 N (%)				
> 200	207 (52.4)	39 (35.8)	128 (58.5)	40 (59.7)
≤ 200	124 (31.4)	12 (11.0)	88 (40.2)	24 (35.8)
Missing	64 (16.2)	58 (53.2)	3 (1.4)	3 (4.5)
Recent CD4[€]; Median (SD)	392.0 (271.4)	392 (287.2)	399.5 (278.1)	381.0 (235.0)
Recent CD4 N (%)				
> 200	268 (67.9)	44 (40.4)	172 (78.5)	52 (77.6)
≤ 200	64 (16.2)	7 (6.4)	44 (20.1)	13 (19.4)
Missing	63 (15.9)	58 (53.2)	3 (1.4)	2 (3.0)

* Defined as no recorded CD4+ T lymphocyte or HIV viral load lab data or most recent lab data within 3 months of HIV diagnosis date.

** Defined as >180 days between any 2 lab dates

^ MSM, men who have sex with men

§ SD, Standard Deviation

° Defined as time from HIV diagnosis date to April 21, 2014

£ First CD4 T lymphocyte lab count recorded in post diagnosis

€ Most recent CD4 T lymphocyte lab count recorded

Table 2. Demographic and Clinical Variables Associated with Having Never Engaged in HIV Care,* N=395

Variable	Bivariate PR (95% CI)¥	p - value	Multivariable PR (95% CI)◆	p - value
Race				
White	Referent			
Black	0.97 (0.57, 1.66)	0.9236		
Hispanic	0.99 (0.44, 2.23)	0.9858		
Other	0.83 (0.31, 2.26)	0.7173		
Gender				
Male	Referent			
Female	0.69 (0.38, 1.25)	0.2016		
Transgender	1.89 (1.24, 2.89)	n/a		
Housing				
Not Homeless	Referent		Referent	
Homeless	1.55 (1.13, 2.13)	0.0073	1.63 (1.21, 2.21)	0.0018
Age				
> 35	Referent		Referent	
26-35	1.58 (1.11, 2.24)	0.0002	1.21 (0.84, 1.76)	0.3166
18-25	3.13 (2.15, 4.55)	0.0152	2.34 (1.46, 3.77)	0.0066
Risk				
Other	Referent			
MSM^	0.87 (0.62, 1.22)	0.4180		
Diagnosis Time°				
> 10 yrs	Referent		Referent	
≤ 10 yrs	3.82 (2.40, 6.09)	<0.0001	3.17 (1.95, 5.15)	<0.0001

* Never engaged in care defined as no recorded CD4+ T lymphocyte or HIV viral load lab data or most recent lab data within 3 months of HIV diagnosis date.

¥ Prevalence Ratio with associated 95% confidence interval

◆ Only variables significant at a level of 0.05 in full bivariate analysis were included in final multivariable model

^ MSM, men who have sex with men

° Defined as time from HIV diagnosis date to April 21, 2014

Table 3. Demographic and Clinical Variables Associated with Having at Least 1 Gap in HIV Care Among Individuals Who Had Ever Engaged in Care, N=286**

Variable	Bivariate PR (95% CI)¥	p - value	Multivariate PR (95% CI)◆	p - value
Race				
White	Referent			
Black	1.21 (0.91, 1.61)	0.1277		
Hispanic	1.38 (1.00, 1.91)	0.0800		
Other	1.08 (0.68, 1.70)	0.7569		
Gender				
Male	Referent			
Female	1.02 (0.86, 1.22)	0.8231		
Transgender	0.66 (0.46, 0.94)	n/a		
Housing				
Not Homeless	Referent			
Homeless	1.01 (0.88, 1.15)	0.9162		
Age				
> 35	Referent		Referent	
26-35	0.80 (0.65, 0.98)	0.0122	0.85 (0.69, 1.04)	0.0671
18-25	1.00 (0.64, 1.57)	0.9921	1.05 (0.72, 1.54)	0.8155
Risk				
Other	Referent			
MSM^	0.96 (0.84, 1.10)	0.5605		
Diagnosis Time°				
> 10 yrs	Referent		Referent	
≤ 10 yrs	0.87 (0.76, 0.99)	0.0357	0.91 (0.79, 1.04)	0.1647
Recent CD4€				
> 200	Referent			
≤ 200	1.01 (0.86, 1.18)	0.9482		

** Defined as >180 days between any 2 lab dates

¥ Prevalence Ratio with associated 95% confidence interval

◆ Only variables significant at a level of 0.05 in full bivariate analysis were included in final multivariable model

^ MSM, men who have sex with men

° Defined as time from HIV diagnosis date to April 21, 2014

€ Most recent CD4 T lymphocyte lab count recorded

Chapter III: Summary, Implications and Future Directions

This study was one of the first to investigate the utility of the development of an HIE between the Georgia Department of Public Health and a local Atlanta clinic. This study is also one of the first to examine factors associated with engagement in care for patients identified via a HIE. Findings from this study indicate that nearly 5% of PLWHA who are out of care in Georgia could potentially be identified via a HIE with a single Atlanta community clinic. These findings suggest that HIE is a feasible means for the Georgia Department of Public Health and community health centers to utilize surveillance data for the identification of patients out of care for HIV. This study also found that younger patients (<35 years) and patients who are homeless are more likely to have never engaged in care. These findings suggest that interventions targeting the needs of these populations are needed to help engage and retain these patients in care.

Future Directions

Future directions for this study would include narrowing the date range of the SJMC clinical encounters in order to identify patients seen at SJMC during their out of care window. Narrowing to this date range would provide an indication of how many patients might have been identified at the point of care had an HIE with provider alerts been in place. There is also a need to examine additional clinical factors that may be associated with engagement such as chief complaint at presentation or the SJMC clinic specialty type where patients were seen. Prior research has demonstrated that patients who have a overall feeling of well being and/or higher CD4 counts are less likely to engage in care (17, 49, 50). However, we know of no studies that have examined the types of clinics or clinic complaints associated with out of care patients who present for

non-HIV care. This information could inform both the development of a HIE and better equip clinics and providers to re-engage patients in care.

Next steps could also include collection of other socio-demographic information that could inform factors relating to engagement in care such as insurance status, income, transportation, and education level. These variables are included in the GDPH eHARS database reports but are rarely populated. Some level of information may be collected at SJMC and other clinics but were not provided for this study.

This study also utilized stringent matching criteria for identifying individuals considered out of care who also presented at SJMC. Patients' first and last name, sex, date of birth and social security number where available were required to match exactly in order to be considered effectively identified and included for further analysis. Next steps could include using less stringent criteria for matching whereby patients could be considered matched on 2 of 4 or 3 of 4 of these criteria. This could increase the sensitivity of the HIE particularly in instances where patients have used nicknames/aliases or alternative social security numbers for clinic paperwork.

Public Health Implications

Successful development and utilization of an alerts based HIE between the GDPH and health providers has great public health implications for the improvement of the proportion of Georgians on the HIV Care Continuum. Timely identification of out of care PLWHA offers an invaluable opportunity for engagement or re-engagement into care. Although this study identified slightly less than 5% of the total out of care population in Georgia, a more substantial number of individuals could potentially be identified if HIE were to exist at large volume primary care community clinics such as

those associated with the Grady Health System and Atlanta Medical Center in the Atlanta metro area. Expansion of HIE to other Georgia health systems such as Piedmont and Dekalb Medical Centers would also extend the potential for identifying out of care patients.

Once out of care patients are identified, health centers and providers need to develop structures and programming to engage those patients in care. Future research needs to be performed to investigate effective strategies for engaging individuals who have never sought care and for re-engaging individuals who have dropped out of care. A study of care engagement by Fleishman et al suggests that these populations may have different needs in addressing their lack of care (17).

This study also investigated the factors that may contribute to a person's engagement status. Our study identified younger patients and homeless patients as those more likely to have never engaged in care. These populations could potentially benefit from structural and systemic interventions such as appointment tracking, extended clinic hours and utilization of community-based organizations. These interventions, along with the use care coordination and patient navigators, have been shown to improve retention of PLWHA (20, 51). These strategies combined with HIE could effectively improve out of care patients' engagement in care in Georgia.

The development and implementation of a successful HIE has positive implications not only for individuals but also for the healthcare system as a whole. Increased retention in care results in increased access to HAART, improved adherence, improved survival and lower health care costs from hospitalization and emergency department visits (20). In addition data collected on the number of persons identified via

HIE could also be used to advocate for funding for increased HIV related community based organizations or HIV care services and clinics in the area. Implementation of HIE in clinics such as SJMC, which predominantly serves Blacks and other minorities could help decrease the disparity for retention in care among those populations. Successful re-engagement in care of those identified has the potential to help decrease disparities in viral load suppression and even survival rates (52).

Appendix A: SAS Code Analysis

```
/*************************************************************
Code: Georgia Out of Care Data set
Author: Eugene Pennisi, Katrece Outlaw
Purpose: Create data set of Georgians out of care as of 4/21/2014
      (from person data as of 11/2014)
Date:   December 1, 2014
Location: H:\Share Drive\DM\EP\HIE\Katrece
Last Modified:

*************************************************************/


/*
Out-of-Care Common Criteria

Include Those Who Meet The Following Criteria:
1.    Meets ALL of the Following Criteria In eHARS
      a.    HIV Diagnosis Code Of 1 Or 2
      b.    Status Flag Of A, W Or R
      c.    HIV Diagnosis Date < Today Minus 455 Days
      d.    "Adult" (Age = 18 Years)
      e.    Vital Status Of "Alive"
      f.    Current Residence Of "GA"
      g.    Has Populated First, Last, DOB, SSN
      h.    =1 Lab Document (Code 004) Indicating WB+ or VL = 401
      i.    No CD4 Or Quantitative VL Done After Today Minus 455 Days
      j.    <= 7 Years Between Most Recent Lab (Any Type) and Today
2.    No Contravening Evidence From CW, SendSS
      a.    Current Address Not In GA or
      b.    =1 CD4 Or Quantitative VL Done After Today Minus 455 Days
or
      c.    Prescribed ART After Today Minus 455 Days (CW)
3.    No Evidence of Denial of Consent to Share HIV Data

No Longer in Care Subset
(HAP message #1)

1.    Meets Above Out-of-Care Common Criteria
2.    Evidence of Prior In-Care Status
      a.    CD4 or Quantitative VL Ever (eHARS, CW, SendSS) or
      b.    Prescribed ART Ever (CW)

Never in Care Subset
(HAP message #2)

1.    Meets Above Out-of-Care Common Criteria
2.    No Evidence of Prior In-Care Status
      a.    No Laboratory Evidence of CD4 or Quantitative VL Ever
(eHARS, CW, SendSS) or
      b.    No Evidence of Prescribed ART Ever (CW)
*/

dm log 'clear';
dm output 'clear';
```

```
options nocenter nodate nonumber nofmterr mprint mlogic symbolgen;

libname per "H:\Share Drive\ehars datasets\PERSON_201411";
libname doc "H:\Share Drive\ehars datasets\DOCUMENT_201411";
libname hie 'H:\Share Drive\DM\EP\HIE\Katrece';

%let today_is=04/21/2014;
%let window=455;

/*
1.    Meets ALL of the Following Criteria In eHARS
      a.    HIV Diagnosis Code Of 1 Or 2
      b.    Status Flag Of A, W Or R
      c.    HIV Diagnosis Date < Today Minus 455 Days
      d.    "Adult" (Age = 18 Years)
      e.    Vital Status Of "Alive"
      f.    Current Residence Of "GA"
      g.    Has Populated First, Last, DOB, SSN
      h.    >=1 Lab Document (Code 004) Indicating WB+ or VL = 401
      i.    No CD4 Or Quantitative VL Done After Today Minus 455 Days
      j.    = 7 Years Between Most Recent Lab (Any Type) and Today
*/
data cases (drop=m d y);
      length _hiv_aids_dx_dt 8. ssn4 $4;
      set per.person;
      if stateno ne ' ';
/*a.  HIV Diagnosis Code Of 1 Or 2*/
      if hiv_categ in ('1','2');

/*b.  Status Flag Of A, W Or R*/
      if status_flag in ('A','W','R');

/*c.  HIV Diagnosis Date < Today Minus 455 Days */
      if hiv_aids_dx_dt not in ('........',' ') and
            substr(hiv_aids_dx_dt,1,4) ne '....';
      if 1980 le int(substr(hiv_aids_dx_dt,1,4));

      *Define _hiv_aids_dx_dt;
      if substr(hiv_aids_dx_dt,5,2) eq '..' then
            m=12;
            else m=int(substr(hiv_aids_dx_dt,5,2));
      if substr(hiv_aids_dx_dt,7,2) eq '..' then
            d=28;
            else d=int(substr(hiv_aids_dx_dt,7,2));
      y=int(substr(hiv_aids_dx_dt,1,4));
      _hiv_aids_dx_dt=mdy(m,d,y);

      if _hiv_aids_dx_dt lt (input("&today_is.",anydtdte10.) -
&window.);

/*d.  "Adult" (Age = 18 Years)*/
      if int(cur_age) ge 18;

/*    e.    Vital Status Of "Alive"*/
      if vital_status='1';
```

```
/*PROPOSED CHANGE*/
/*    f.    Current Residence Of "GA"*/
      if cur_state_cd='GA';

/*    change*/
/*    if cur_state_cd in ('GA',' ');*/

/*g.  Has Populated First, Last, DOB, SSN*/
      if dob ne ' ';
      if index(dob,'.')=0;   *Could loosen this up?;
      if index(ssn,'-') not in (4,7) then ssn='.........';
            else ssn=compress(ssn,'-');
      if ssn='999999999' then ssn='.........';
            else if substr(ssn,1,1)='9' then substr(ssn,1,5)='.....';
      if substr(ssn,1,3)='000' or substr(ssn,4,2)='00' then
substr(ssn,1,5)='.....';
      if substr(ssn,6,4)='0000' then ssn='.........';
      if substr(ssn,6,4) ne '....' then ssn4=substr(ssn,6,4);
            else ssn4='....';
      if last_name ne ' ';
      if first_name ne ' ';

/*    PROPOSED CHANGE*/
      if length(first_name) ge 2 and length(last_name) ge 2;
      if notalpha(compress(last_name))=0 and
notalpha(compress(first_name))=0;

/*    change*/
/*    if length(last_name) ge 2;*/
run;

/*h.  =1 Lab Document (Code 004) Indicating WB+ or VL >= 401*/
proc sql;
      create table cases2 as
      select distinct
            a.*
      from
            cases as a,
            doc.document as b,
            doc.lab as c
      where
            a.ehars_uid=b.ehars_uid and
            b.document_uid=c.document_uid and
            b.document_type_cd='004' and
            b.status_flag='A' and
            c.sample_dt ne ' ' and substr(c.sample_dt,1,4) ne '....'
and
            (2003 le input(substr(c.sample_dt,1,4),8.)) and
            (input("&today_is.",anydtdte10.) -
input(substr(c.sample_dt,5,2)||"/"

      ||substr(c.sample_dt,7,2)||"/"||substr(c.sample_dt,1,4),anydtdte1
0.) > &window.) and
            (
                  (c.lab_test_cd in ('EC-006','EC-007') and
result='POS')
                  or
```

```
                  (c.lab_test_cd in ('EC-014','EC-015') and
                        input(result,8.)>400 and result_units='C/ML'
and result_interpretation in ('=','>'))
                  or
                  (c.lab_test_cd ='EC-008' and result='POS')
            );
quit;

data cases3;
      length _recent_cd4_vl 8. _cd4_recent_cnt_pct_dt 8.;
      length never_in_care $1;
      set cases2;
/*i.  No CD4 Or Quantitative VL Done After Today Minus 455 Days*/
      if cd4_recent_cnt_pct_dt ne ' ' and
            substr(cd4_recent_cnt_pct_dt,1,4) ne '....' then
            do;
                  if substr(cd4_recent_cnt_pct_dt,5,2) eq '..' then
                        m=12;
                        else m=int(substr(cd4_recent_cnt_pct_dt,5,2));
                  if substr(cd4_recent_cnt_pct_dt,7,2) eq '..' then
                        d=28;
                        else d=int(substr(cd4_recent_cnt_pct_dt,7,2));
                  y=int(substr(cd4_recent_cnt_pct_dt,1,4));
                  _cd4_recent_cnt_pct_dt=mdy(m,d,y);
                  drop m d y;
            end;
      if vl_recent_dt ne ' ' and
            substr(vl_recent_dt,1,4) ne '....' then
            do;
                  if substr(vl_recent_dt,5,2) eq '..' then
                        m=12;
                        else m=int(substr(vl_recent_dt,5,2));
                  if substr(vl_recent_dt,7,2) eq '..' then
                        d=28;
                        else d=int(substr(vl_recent_dt,7,2));
                  y=int(substr(vl_recent_dt,1,4));
                  _vl_recent_dt=mdy(m,d,y);
                  drop m d y;
            end;
      if _cd4_recent_cnt_pct_dt ne . or _vl_recent_dt ne . then
            do;

      _recent_cd4_vl=max(_cd4_recent_cnt_pct_dt,_vl_recent_dt);
                  never_in_care='N';
            end;
      if _recent_cd4_vl ge (input("&today_is.",anydtdte10.)-&window.)
      then delete;
      else if _recent_cd4_vl = . then never_in_care='Y';
run;

/*    j.<= 7 Years Between Most Recent Lab (Any Type) and Today*/
data cases4;
      length birth_dt $10;
      set cases3;
      if test_recent_dt ne ' ' and substr(test_recent_dt,1,4) ne
'....';
```

```
      if year(input("&today_is.",anydtdte10.)) -
input(substr(test_recent_dt,1,4),8.) le 7;
      birth_dt=substr(dob,5,2)||'/'||substr(dob,7,2)||'/'||substr(dob,1
,4);
      recent_cd4_vl=put(_recent_cd4_vl,mmddyy10.);
run;

proc sort data=cases4 out=cases_export
      (keep=never_in_care stateno last_name middle_name first_name
birth_dt ssn);
      by never_in_care stateno;
run;

data cases_export;
      retain never_in_care stateno last_name middle_name first_name
birth_dt ssn;
      set cases_export;
run;

data hie.stjoe_outofcare;
      set cases4;
run;

*OOC List 8350 obs;
/*PROGRAM ENDS HERE*/

/*************************************************************
Code: SJMC Matches data set
Author: Katrece Outlaw and Eugene Pennisi
Purpose: Create data set of matches between STJ encounters, eHARS OOC
list
      (from person data as of 11/2014) and eHARS lab document data as
of 11/2014.

Date:   December 3, 2014
Location: H:\Share Drive\DM\EP\HIE\Katrece
Last Modified:

*************************************************************/

dm log 'clear';
dm output 'clear';

options nocenter nodate nonumber nofmterr mprint mlogic symbolgen;

libname per "H:\Share Drive\ehars datasets\PERSON_201411";
libname doc "H:\Share Drive\ehars datasets\DOCUMENT_201411";
libname stjoe 'H:\Share Drive\DM\EP\HIE\Katrece\data sets';

libname stj 'H:\Share Drive\DM\EP\HIE\st joe';
libname rev 'H:\Share Drive\DM\Core - SAS Code\EP';

%let today_is=04/21/2014;
%let window=455;

/*********Matching STJ Encounters to all  EHARS***************/
PROC IMPORT OUT= WORK.top1000
```

```
            DATAFILE= "H:\Share Drive\DM\EP\HIE\Top1000.xls"
            DBMS=EXCEL REPLACE;
     RANGE="top1000$";
     GETNAMES=YES;
     MIXED=YES;
     SCANTEXT=YES;
     USEDATE=YES;
     SCANTIME=YES;
RUN;

data top100 (keep=name);
      set top1000 (keep=Table_with_row_headings_in_colum f2);
      if f2 le 100;
      if f2="rank" then delete;
      rename Table_with_row_headings_in_colum=name;
      rename f2=rank;
run;

data stjoe;
      set stj.st_joe_encounters;
      array vars(*) _character_;
      do i=1 to dim(vars);
      vars(i)=upcase(vars(i));
      vars(i)=strip(vars(i));
      vars(i)=compbl(vars(i));
      vars(i)=compress(vars(i),,'wk');
      end;
      drop i;
run;

*cleaning stj encounter data in preparation for match;
data stjoe2 (rename=(_ssn=ssn _dob=dob));
      length ssn4 $4 _ssn $9;
      set stjoe;
      if dob =. then delete;
            else _dob=put(dob,yymmdds10.);
      drop dob;
      last_name=tranwrd(last_name,'3rd','III');
      if anydigit(last_name) ne 0 then delete;
      if anydigit(first_name) ne 0 then delete;
      if last_name='DO NOT DOUBLE BOOK' then delete;
      if first_name='DOUBLE BOOK' then delete;
      if index(last_name,'DO NOT') ne 0 then delete;
      _ssn=left(compress(put(ssn,9.)));
      if _ssn='.' then _ssn=' ';
      if _ssn in
('111111111','222222222','333333333','444444444','555555555','666666666
',
            '777777777','888888888','999999999') then _ssn=' ';
      if substr(_ssn,1,3)='999' then _ssn=' ';
      if length(_ssn) ge 4 then ssn4=substr(right(_ssn),6,4);
      if length(compress(_ssn)) ne 9 then _ssn=' ';
      drop ssn;
run;

*cleaning ehars data in preparation for match;
data ehars (rename=(_ssn=ssn _dob=dob));
```

```
      length ssn4 $4 _ssn $9;
      set per.person (keep=stateno ssn dob race status_flag first_name
      last_name hiv_categ);
      if stateno ne ' ';
      if hiv_categ in ('1','2');
      if status_flag in ('A','W','R','E');
      if dob=' ' then delete;
            else _dob=put(input(dob,anydtdte10.),yymmdds10.);
      drop dob;
      _ssn=compress(ssn,' -');
      if _ssn in
('111111111','222222222','333333333','444444444','555555555','666666666
','777777777','888888888','999999999') then _ssn=' ';
      if _ssn ne ' ' and substr(_ssn,1,3)='000' then _ssn=' ';
      if _ssn ne ' ' and index(substr(_ssn,6,4),'.') ne 0 then _ssn='
      ';
            else if _ssn ne ' ' then ssn4=substr(_ssn,6,4);
      if _ssn ne ' ' and index(_ssn,'.') ne 0 then _ssn=' ';
      drop ssn;
      if last_name=' ' or first_name=' ' then delete;
      last_name=compbl(compress(last_name,,'apsk'));
      first_name=compbl(compress(first_name,,'apsk'));
      if length(first_name) ge 4 then fn4=substr(first_name,1,4);
      if length(last_name) ge 4 then ln4=substr(last_name,1,4);
run;

**running match with cleaned data using ehars data set as of 11/2014,
incorporating homelessness and last encounter
date from stj encounters data;

*match based on f/l name, dob, ssn, ssn4;
proc sql;
      create table match as
      select distinct
            strip(b.stateno) as stateno,
            strip(a.race) as a_race,
            strip(b.race) as b_race,
            strip(a.ssn) as a_ssn,
            strip(b.ssn) as b_ssn,
            a.dob as a_dob,
            b.dob as b_dob,
            strip(a.last_name) as a_last_name,
            strip(b.last_name) as b_last_name,
            strip(a.first_name) as a_first_name,
            strip(b.first_name) as b_first_name,
            a.homelessstatus as homelessstatus,
            a.last_enc_date as last_enc_date

      from
            (select distinct last_name, first_name, dob, ssn, ssn4,
            race,
                  homelessstatus, last_enc_date from stjoe2) as a,
            ehars (keep= stateno race last_name first_name dob ssn
            ssn4) as b
      where
            (
                  a.last_name=b.last_name and
```

```
                  a.dob=b.dob and
                  a.ssn4=b.ssn4 and
                  a.ssn4 ne ' '
            ) or
            (
                  (
                        a.last_name=b.last_name and
                        a.last_name not in (select * from top100)
                  ) and
                  a.first_name=b.first_name and
                  a.dob=b.dob
            )
;
quit;

*looking for duplicates in dataset;
proc sort data=match dupout=mydup nodupkey; by stateno; run;

proc print data=mydup noobs;
var stateno a_last_name b_last_name a_first_name b_first_name a_dob
b_dob a_ssn b_ssn last_enc_date; run;

*idk what this does;
data match2;
      set match;
      last_name2=last_name;
      substr(a_last_name,1,1)="";
      a_last_name=strip(a_last_name);
run;

*look at match;
proc contents data=match2; run;

*set permanent data set;
data stjoe.match_angela;
      set match2;
run;


/********matching stj matches from all ehars to OOC LIST******/

*renaming permanent datasets to create temp ones;
data stjoe_ehars;
      set stjoe.match_angela;
run;

data ehars_ooc;
      set stjoe.stjoe_outofcare;
run;

*matching to OOC list as of 4/21/2014 using 11/2014 ehars person data;
proc sql;
      create table stjoe_ooc as
      select
            a.*,c.*
      from
```

```
            stjoe_ehars (keep=stateno homelessstatus last_enc_date) as
            a,
            ehars_ooc as b,
            per.person as c
            where
            compress(a.stateno)=compress(b.stateno) and
            compress(a.stateno)=compress(c.stateno);
quit;


/**************Looking at matched data******************/
proc contents data=stjoe_ooc; run;

*looking for duplicates;
proc sort data=stjoe_ooc out=test nodupkey; by stateno; run;

proc sort data=stjoe_ooc dupout=dups nodupkey; by stateno; run;

proc print data=dups noobs; var stateno last_name first_name
last_enc_date; run;

*create permanent data set of matches;
data stjoe.stjoe_ooc2;
      set  stjoe_ooc;
run;

/*********create data set with only desired variables*******/
data matchvar;

keep StateNo SSN dob last_name first_name
aids_age_yrs hiv_aids_age_yrs cur_age homelessstatus last_enc_date
cd4_recent_cnt_value cd4_recent_cnt_dt vl_recent_value vl_recent_dt
birth_sex current_gender ethnicity1 Race disease_categ_dx
trans_categ aids_insurance aids_cdc rsx_county_name rsx_state_cd
cd4_vl_first_hiv_dt
cd4_first_hiv_value document_uid hiv_aids_dx_dt ;

set stjoe.stjoe_ooc2;

run;

/****linking STJ-OOC matched dataset to ehars docs***********/

*cleaning document data for linkage;
*restrict lab data to CD4/VL cnt/pct as of 2006 for better accuracy of
data;
data care;
      set doc.lab;
      if lab_test_cd in ('EC-014','EC-015','EC-016','EC-017');
      if sample_dt not in (' ','........');
      if substr(sample_dt,1,4) ne '....';
      if 2006 le input(substr(sample_dt,1,4),8.) le 2013;
run;

*matching to lab doc data;
proc sql;
      create table docmatches as
```

```
      select
            a.*, b.*
      from
            matchvar as a left outer join
            care(keep=document_uid sample_dt result lab_test_type
result_units) as b
      on
            a.document_uid=b.document_uid and
            input(substr(a.hiv_aids_dx_dt,1,4),4.) le
input(substr(b.sample_dt,1,4),4.);
quit;

*create permanent data set;
data stjoe.stjdocmatches;
set docmatches;
run;
```

```
/*************************************************************************
Code: SJMC Matches cleaning and variable creation
Author: Katrece Outlaw and Eugene Pennisi
Purpose: Modify data set of STJ-OOC matches for analysis

Date:   December 3, 2014
Location: H:\Share Drive\DM\EP\HIE\Katrece
Last Modified:

*************************************************************************/

dm log 'clear';
dm output 'clear';

options nocenter nodate nonumber nofmterr mprint mlogic symbolgen;

libname stjoe 'H:\Share Drive\DM\EP\HIE\Katrece\data sets';
libname stj 'H:\Share Drive\DM\EP\HIE\st joe';

%let today_is=04/21/2014;
%let window=455;


data docmatches2;
set stjoe.stjdocmatches;

if current_gender='U' then curr_gender= ' ';
else curr_gender=current_gender;

*creating numeric variables for cnts and ages;
cd4_recent_cnt= cd4_recent_cnt_value*1;
vl_recent_cnt= vl_recent_value*1;
cur_age1= cur_age*1;
hiv_aids_age= hiv_aids_age_yrs*1;
cd4_first_value=cd4_first_hiv_value*1;

*making all missing date days to 15;
if substr(vl_recent_dt,7,2)='..' then
vl_recent_dt2=substr(vl_recent_dt,1,6)||'15';
else vl_recent_dt2=vl_recent_dt;

if substr(cd4_recent_cnt_dt,7,2)='..' then
cd4_recent_dt=substr(cd4_recent_cnt_dt,1,6)||'15';
else cd4_recent_dt=cd4_recent_cnt_dt;

if substr(cd4_vl_first_hiv_dt,7,2)='..' then
cd4_vl_first_dt=substr(cd4_vl_first_hiv_dt,1,6)||'15';
else cd4_vl_first_dt=cd4_vl_first_hiv_dt;

if substr(sample_dt,7,2)='..' then
sample_dt2=substr(sample_dt,1,6)||'15';
else sample_dt2=sample_dt;

*convert to date format;
_vl_recent_dt2= input(vl_recent_dt2,anydtdte8.);
_cd4_recent_dt= input(cd4_recent_dt,anydtdte8.);
_cd4_vl_first_dt=input(cd4_vl_first_dt,anydtdte8.);
```

```
_sample_dt2= input(sample_dt2,anydtdte8.);

*want to create variable 'most recent' where most recent of cd4 or
viral load;
if _cd4_recent_dt = . and _vl_recent_dt2= . then most_recent= .;
else if _cd4_recent_dt ge _vl_recent_dt2 then
most_recent=_cd4_recent_dt;
else most_recent=_vl_recent_dt2;


*creating nevercare variable - where first lab date and most recent
data w/in same month;
if most_recent=. or _cd4_vl_first_dt=. then nevercare=1;
else if most_recent - _cd4_vl_first_dt lt 91 then nevercare=1;
else if year(most_recent) lt 2006 then nevercare=.;
else nevercare=0;

format _cd4_vl_first_dt _vl_recent_dt2 _cd4_recent_dt most_recent
yymmdd10.;
format _sample_dt2 yymmdd10.;

run;

proc contents data=docmatches2; run;

*check creation of new variables;
proc freq data=docmatches2; tables
_cd4_recent_dt*_vl_recent_dt2*most_recent/ list missing;
run;

proc freq data=docmatches2; tables
_cd4_vl_first_dt*most_recent*nevercare/ list missing; /*where
nevercare=1*/; run;

proc freq data=docmatches2; tables current_gender*curr_gender/ list
missing; run;


********creation of more variables******;
proc sort data=docmatches2 out=docmatch_day nodupkey;
by document_uid _sample_dt2; run;

data docmatch_day2;
      set docmatch_day;
      by document_uid _sample_dt2;

      lag_day=(_sample_dt2-lag(_sample_dt2));
      lag_mo=(_sample_dt2-lag(_sample_dt2))/30;
      lag2_day=(_sample_dt2-lag2(_sample_dt2));


      if first.document_uid then lag_day=.;
      if first.document_uid then lag_mo=.;
      if first.document_uid then lag2_day=.;
      if lag2_day lt 0 then lag2_day=.;

      ID=input(compress(document_uid,,'dk'),21.)+101013030;
```

```
      T1days=(mdy(4, 21, 2014)-most_recent);
      T1yrs=(mdy(4, 21, 2014)-most_recent)/365.25;
      T1exd=((mdy(4, 21, 2014)-most_recent)-455);
      T1exy=((mdy(4, 21, 2014)-most_recent)-455)/365.25;

      *creating age category variables;
      if cur_age1= . then age_cat=.;
      else if cur_age1 ge 18 and cur_age1 le 29 then age_cat=1;
      else if cur_age1 ge 30 and cur_age1 le 39 then age_cat=2;
      else if cur_age1 ge 40 and cur_age1 le 49 then age_cat=3;
      else if cur_age1 ge 50 and cur_age1 le 59 then age_cat=4;
      else if cur_age1 ge 60 then age_cat=5;

      *creating category variable for age at diagnosis;
      if hiv_aids_age= . then hiv_age_cat=.;
      else if hiv_aids_age ge 14 and hiv_aids_age le 29 then
hiv_age_cat=1;
      else if hiv_aids_age ge 30 and hiv_aids_age le 39 then
hiv_age_cat=2;
      else if hiv_aids_age ge 40 and hiv_aids_age le 49 then
hiv_age_cat=3;
      else if hiv_aids_age ge 50 and hiv_aids_age le 59 then
hiv_age_cat=4;
      else if hiv_aids_age ge 60 then hiv_age_cat=5;

      *creating simplified housing status variable;
      if homelessstatus = ' ' then hsing_status= .;
      else if homelessstatus = 'NOT HOMELESS' then hsing_status = 1;
      else if homelessstatus = 'SHELTER (HL)' or homelessstatus
      ='BOARDING HOUSE/HOTEL (HL)'
            then hsing_status = 2;
      else if homelessstatus = 'STREET (HL)' or homelessstatus
      ='TRANSITIONAL (HL)'
            then hsing_status = 3;
      else if homelessstatus = 'DOUBLING UP / FRIENDS/RELATIVES (HL)'
            then hsing_status = 4;
      else if homelessstatus = 'HALFWAY HOUSE / SUBSTANCE ABUSE
      FACILITY (HL)'
            then hsing_status=5;
      else if homelessstatus= 'JAIL OR CORRECTIONAL FACILITY (HL)'
      then hsing_status=6;
      else if homelessstatus= 'HOSPITAL (HL)'
            then hsing_status=7;
      else if homelessstatus= 'DOMESTIC VIOLENCE FACILITY'
      then hsing_status = 8;
      else hsing_status= 9;

      if cd4_first_value=. then cd4_first_cat=.;
      else if cd4_first_value lt 200 then cd4_first_cat=1;
      else if cd4_first_value ge 200 and cd4_first_value lt 350
      then cd4_first_cat=2;
      else if cd4_first_value ge 350 and cd4_first_value lt 500
      then cd4_first_cat=3;
      else if cd4_first_value ge 500 then cd4_first_cat=4;

      if cd4_recent_cnt=. then cd4_recent_cat=.;
```

```
      else if cd4_recent_cnt le 200 then cd4_recent_cat=1;
      else if cd4_recent_cnt gt 200 and cd4_recent_cnt lt 350
      then cd4_recent_cat=2;
      else if cd4_recent_cnt gt 350 and cd4_recent_cnt lt 500
      then cd4_recent_cat=3;
      else if cd4_recent_cnt ge 500 then cd4_recent_cat=4;

      last_enc_yr= year(last_enc_date);

run;

*create permanent data set;
data stjoe.matchvar_dup;
      set docmatch_day2;
run;


*****create de-identified version of matchvar_dup;
data de_id;
set stjoe.matchvar_dup (drop= last_name first_name ssn dob stateno
document_uid _cd4_recent_dt
 _cd4_vl_first_dt _sample_dt2 _vl_recent_dt2  cd4_recent_cnt_dt
cd4_recent_dt
cd4_vl_first_dt   cd4_vl_first_hiv_dt last_enc_date sample_dt
      sample_dt2 disease_categ_dx
most_recent vl_recent_dt  vl_recent_dt2  rsx_county_name  rsx_state_cd
hiv_aids_dx_dt);

run;


/*******creation of lagmax variable, will de-dup dataset*********/

*first sort by stateno;
proc sort data=docmatch_day2 out=docmatch_day3; by stateno; run;

data docmatch_max;
  set docmatch_day3;
  by stateno;

  lag_max = max(lag_max,lag_day );

  if last.stateno then do;
      output docmatch_max;
      lag_max=.;
  end;

  retain lag_max;

run;

proc sort data=docmatch_max out=docsort; by stateno; run;

proc sort data=freqtest out=freq; by stateno; run;
```

```
data docmatchmax_ooc;
      merge docsort freq;
      by stateno;

      rename COUNT= ooc_cnt;
      label ooc_cnt='ooc_cnt';
run;

proc contents data=     docmatchmax_ooc; run;

proc freq data=   docmatchmax_ooc; tables ooc_cnt; run;

proc print data=docmatchmax_ooc noobs; var stateno lag_max ooc_cnt
nevercare;
where ooc_cnt=.;
run;

*create permanent data set de-duped plus ooc;
data stjoe.matchvar_tot_ooc;
      set docmatchmax_ooc;
run;

proc print data= stjoe.matchvar_tot noobs;
var stateno _sample_dt2 lag_max nevercare;
where nevercare=0 and lag_max=.;
run;

*create variable for time since diagnosis (yrs/months);
data docmatchmax_ooc2;
      set stjoe.matchvar_tot_ooc;

      if substr(hiv_aids_dx_dt,5,4)='....' then
hiv_dx_dt=substr(hiv_aids_dx_dt,1,4)||'0115';
      else if substr(hiv_aids_dx_dt,7,2)='..' then
hiv_dx_dt=substr(hiv_aids_dx_dt,1,6)||'15';
      else hiv_dx_dt=hiv_aids_dx_dt;

      _hiv_dx_dt= input(hiv_dx_dt,anydtdte8.);

      format _hiv_dx_dt yymmdd10.;

      dx_time=(mdy(4, 21, 2014)-_hiv_dx_dt)/365.25;

      dx_time_mo=(mdy(4, 21, 2014)-_hiv_dx_dt)/30;

run;

*create de-identified data set of docmatchmax_ooc2;
data de_id2;

set docmatchmax_ooc2 (drop= last_name first_name ssn dob stateno
document_uid _cd4_recent_dt
 _cd4_vl_first_dt _sample_dt2 _vl_recent_dt2  cd4_recent_cnt_dt
cd4_recent_dt
cd4_vl_first_dt   cd4_vl_first_hiv_dt last_enc_date sample_dt
      sample_dt2 disease_categ_dx
```

```
most_recent vl_recent_dt  vl_recent_dt2  rsx_county_name  rsx_state_cd
hiv_aids_dx_dt
hiv_dx_dt _hiv_dx_dt);

run;

proc contents data=de_id2; run;


*create permanent data set;
data stjoe.matchvar_deid2;
set de_id2;
run;
```

```
/*************************************************************************

Author: Katrece Outlaw
Purpose: SJMC Matches variable creation and descriptive stats
Date:   December 3, 2014
Location: H:\Thesis
Last Modified:

*************************************************************************/

dm log 'clear';
dm output 'clear';

options nocenter nodate nonumber nofmterr mprint mlogic symbolgen;

libname stj 'H:\Thesis';
****;
data stj.dedup_thesis_late;
set dedup_thesis;
run;

proc contents data=dedup_thesis; run;


*fixing some variables;
data dedup_thesis;
      set stj.dedup_thesis_late;

      if ooc_cnt=. then ooc_cnt=0;
      if lag_max=. then nevercare=1;

      if race=. then race_cat=' ';
      else if race= 4 then race_cat= 'B';
      else if race=6 then race_cat='W';
      else if race=1 then race_cat='H';
      else race_cat= 'O';

      if hsing_status=. then hse_cat=.;
      else if hsing_status=1 then hse_cat=0;
      else hse_cat=1;

      if trans_categ=. then trans_categ2='     ';
      else if trans_categ= 1 or trans_categ= 3
      then trans_categ2= 'MSM';
      else trans_categ2= 'other';

      if cur_age1= . then age_cat3=.;
      else if cur_age1 ge 18 and cur_age1 le 25 then age_cat3=1;
      else if cur_age1 ge 26 and cur_age1 le 35 then age_cat3=2;
      else if cur_age1 gt 35 then age_cat3=3;

      if hiv_aids_age=. then hiv_age_bin=.;
      else if hiv_aids_age le 25 then hiv_age_bin=1;
      else hiv_age_bin=0;


      if current_gender='  ' then curr_gender2='  ';
```

```
else if current_gender='U' then curr_gender2='  ';
else curr_gender2=current_gender;

if lag_max=. then lag_max1=.;
else if lag_max gt 180 then lag_max1=1;
else lag_max1=0;

if cd4_first_value = . then cd4_first_bin=.;
else if cd4_first_value le 200 then cd4_first_bin=1;
else cd4_first_bin=0;

if cd4_first_value = . then cd4_first_bin1=.;
else if cd4_first_value le 100 then cd4_first_bin1=1;
else cd4_first_bin1=0;

if cd4_first_value = . then cd4_first_cat=.;
else if cd4_first_value lt 100 then cd4_first_cat=1;
else if cd4_first_value ge 100 and cd4_first_value lt 200
then cd4_first_cat=2;
else if cd4_first_value ge 200 and cd4_first_value lt 300
then cd4_first_cat=3;
else if cd4_first_value ge 300 and cd4_first_value lt 400
then cd4_first_cat=4;
else if cd4_first_value ge 400 and cd4_first_value lt 500
then cd4_first_cat=5;
else if cd4_first_value ge 500 then cd4_first_cat=6;

if cd4_recent_cnt=. then cd4_recent_bin=.;
else if cd4_recent_cnt le 200 then cd4_recent_bin=1;
else cd4_recent_bin=0;

if cd4_recent_cnt = . then cd4_recent_cat=.;
else if cd4_recent_cnt lt 100 then cd4_recent_cat=1;
else if cd4_recent_cnt ge 100 and cd4_recent_cnt lt 200
then cd4_recent_cat=2;
else if cd4_recent_cnt ge 200 and cd4_recent_cnt lt 300
then cd4_recent_cat=3;
else if cd4_recent_cnt ge 300 and cd4_recent_cnt lt 400
then cd4_recent_cat=4;
else if cd4_recent_cnt ge 400 and cd4_recent_cnt lt 500
then cd4_recent_cat=5;
else if cd4_recent_cnt ge 500 then cd4_recent_cat=6;

if dx_time=. then dx_time_cat=.;
else if dx_time ge 1 and dx_time le 5 then dx_time_cat=1;
else if dx_time gt 5 and dx_time le 10 then dx_time_cat=2;
else if dx_time gt 10 and dx_time le 15 then dx_time_cat=3;
else if dx_time gt 15 and dx_time le 20 then dx_time_cat=4;
else if dx_time gt 20 and dx_time le 25 then dx_time_cat=5;
else if dx_time gt 25 and dx_time le 30 then dx_time_cat=6;
else dx_time_cat=7;

if dx_time=. then dx_time_cat2=.;
else if dx_time ge 1 and dx_time le 10 then dx_time_cat2=1;
else if dx_time gt 10 and dx_time le 20 then dx_time_cat2=2;
else if dx_time gt 20 and dx_time le 30 then dx_time_cat2=3;
else dx_time_cat2=4;
```

```
      if dx_time=. then dx_time_bin=.;
      if dx_time le 5 then dx_time_bin=1;
      else dx_time_bin = 0;

      if dx_time=. then dx_time_bin2=.;
      if dx_time le 10 then dx_time_bin2=1;
      else dx_time_bin2 = 0;

      if ooc_cnt=. then ooc_cnt_cat=.;
      else if ooc_cnt= 1 then ooc_cnt_cat=1;
      else if ooc_cnt=2 then ooc_cnt_cat=2;
      else if ooc_cnt gt 2 then ooc_cnt_cat=3;


run;

*looking at demographic statistics for variable data;
proc univariate data=dedup_thesis; var cd4_recent_cnt vl_recent_cnt
cur_age1 hiv_aids_age cd4_first_value;
histogram  cd4_recent_cnt vl_recent_cnt cur_age1 hiv_aids_age
cd4_first_value;
run;

proc freq data=dedup_thesis;
tables race_cat curr_gender2 hse_cat age_cat3
      trans_categ2 cd4_first_bin dx_time_bin2;
run;

proc freq data=dedup_thesis;
tables race_cat curr_gender2 hse_cat age_cat3
      trans_categ2 cd4_first_bin dx_time_bin2;
where nevercare=1;
run;

proc freq data=dedup_thesis;
tables race_cat curr_gender2 hse_cat age_cat3
      trans_categ2 cd4_first_bin dx_time_bin2;
where nevercare=0 and lag_max1=1;
run;
*examining for any association between predictor variables;
proc freq data=dedup_thesis;
tables hse_cat*curr_gender2/ fisher;
run;

proc freq data=dedup_thesis;
tables hse_cat*age_cat3/ chisq;
run;

proc format;

value age_cat 1-2="Index"
                        3="Reference";

value yesnoa      . = "Missing"
                        0 = "2 No"
```

```
                        1 = "1 Yes";
run;

*are younger less likely to be homeless?;
proc freq data=dedup_thesis order=formatted;
format hse_cat yesnoa. age_cat3 age_cat.;
tables hse_cat*age_cat3/ riskdiff cmh;
where age_cat3=1 or age_cat=3;
run;

proc freq data=dedup_thesis order=formatted;
format hse_cat yesnoa. age_cat3 age_cat.;
tables hse_cat*age_cat3/ riskdiff cmh;
where age_cat3=2 or age_cat=3;
run;
*OR shows less likely but not significant;

proc freq data=dedup_thesis;
tables age_cat3*hse_cat;
where nevercare=1;
run;

proc freq data=dedup_thesis;
tables dx_time_bin2*age_cat3/ chisq;
run;

proc freq data=dedup_thesis;
tables dx_time_bin2*cd4_first_bin/ chisq;
run;
```

```
/*************************************************************************

Author: Katrece Outlaw
Purpose: SJMC Matches prevalence ratio analysis

Date:   February 6, 2015
Location: H:\Thesis
Last Modified
*************************************************************************/

dm log 'clear';
dm output 'clear';

options nocenter nodate nonumber nofmterr mprint mlogic symbolgen;

libname stj 'H:\Thesis';


/*********************Looking at prevalence ratios *****************/

data dedup_thesisx;
      set dedup_thesis;

      if race_cat=' ' then race_cat2=.;
      else if race_cat='W' then race_cat2=0;
      else if race_cat='B' then race_cat2=1;
      else if race_cat='H' then race_cat2=2;
      else race_cat2=3;

      if curr_gender2=' ' then curr_gender3=.;
      else if curr_gender2='M' then curr_gender3=0;
      else if curr_gender2='F' then curr_gender3=1;
      else curr_gender3=2;

      if trans_categ2='     ' then trans_categ3=.;
      else if trans_categ2= 'other' then trans_categ3=0;
      else trans_categ3=1;

run;

proc freq data=dedup_thesisx; tables curr_gender2*curr_gender3/ list
missing; run;

proc rlogist data=dedup_thesis design=srs;
      class   age_cat22 dx_time_bin;
      reflev  age_cat22=0 dx_time_bin=0;
      model nevercare= age_cat22 dx_time_bin;
      predmarg age_cat22(0) dx_time_bin(0) /adjrr;
run;
*age no longer significant;



*try stratified;
proc sort data=dedup_thesisx out=bivar_the; by hse_cat; run;

proc rlogist data=bivar_the design=srs;
```

```
      class age_cat3 hse_cat;
      reflev age_cat3=3 hse_cat=0;
      model  nevercare= age_cat3 hse_cat age_cat3*hse_cat;
      predmarg age_cat3(3)*hse_cat(0)/adjrr;
      predmarg age_cat3(3)*hse_cat(1)/adjrr;
run;

*bivariate PRs;
proc rlogist data=dedup_thesisx;
      class   age_cat3 ;
      reflev  age_cat3=3;
      model nevercare= age_cat3;
      predmarg age_cat3(3) /adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
      class   dx_time_bin2;
      reflev   dx_time_bin2=0;
      model nevercare= dx_time_bin2;
      predmarg dx_time_bin2(0) /adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
      class  hse_cat ;
      reflev  hse_cat=0;
      model nevercare= hse_cat;
      predmarg hse_cat(0) /adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
      class  cd4_first_bin;
      reflev cd4_first_bin=0;
      model nevercare= cd4_first_bin;
      predmarg cd4_first_bin(0) /adjrr;
run;


proc rlogist data=dedup_thesisx design=srs;
      class  race_cat2;
      reflev race_cat2=0;
      model nevercare= race_cat2;
      predmarg race_cat2(0) /adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
      class  curr_gender3;
      reflev curr_gender3=0;
      model nevercare= curr_gender3;
      predmarg curr_gender3(0) /adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
      class  trans_categ3;
      reflev trans_categ3=0;
      model nevercare= trans_categ3;
      predmarg trans_categ3(0) /adjrr;
run;
```

```
*try interaction term - please work;
proc rlogist data=dedup_thesis design=srs;
      class   age_cat3 hse_cat;
      reflev  age_cat3=3;
      model nevercare= age_cat3;
      predmarg age_cat3(3) /adjrr;
run;


*FULL MODEL;
proc rlogist data=dedup_thesisx design=srs;
      class  hse_cat age_cat3 cd4_first_bin dx_time_bin2 trans_categ3
curr_gender3 race_cat2;
      reflev cd4_first_bin=0 hse_cat=0 age_cat3=3 dx_time_bin2=0
trans_categ3=0
            curr_gender3=0 race_cat2=0;
      model nevercare= hse_cat age_cat3 cd4_first_bin dx_time_bin2
trans_categ3
            curr_gender3 race_cat2;
      predmarg cd4_first_bin(0) hse_cat(0) age_cat3(3) dx_time_bin2(0)
curr_gender3(0)
                  trans_categ3(0) race_cat2(0)/adjrr;
run;

*trying model with variables suggested via selection;
proc rlogist data=dedup_thesisx design=srs;
      class  hse_cat age_cat3 dx_time_bin2;
      reflev  hse_cat=0 age_cat3=3 dx_time_bin2=0;
      model nevercare= hse_cat age_cat3 dx_time_bin2;
      predmarg  hse_cat(0) age_cat3(3) dx_time_bin2(0)/adjrr;
run;

*trying model with variables suggested by bivariate analysis;
proc rlogist data=dedup_thesisx design=srs;
      class  age_cat3 cd4_first_bin dx_time_bin2 hse_cat;
      reflev cd4_first_bin=0 age_cat3=3 dx_time_bin2=0 hse_cat=0;
      model nevercare=  age_cat3 cd4_first_bin dx_time_bin2 hse_cat;
      predmarg cd4_first_bin(0) age_cat3(3) dx_time_bin2(0)
hse_cat(0)/adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
      class  hse_cat cd4_first_bin1 dx_time_bin2;
      reflev cd4_first_bin1=0 hse_cat=0 dx_time_bin2=0;
      model nevercare= hse_cat cd4_first_bin1 dx_time_bin2;
      predmarg cd4_first_bin1(0) hse_cat(0) dx_time_bin2(0)/adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
      class  hse_cat age_cat3 dx_time_bin2;
      reflev  hse_cat=0 age_cat3=3 dx_time_bin2=0;
      model nevercare= hse_cat age_cat3  dx_time_bin2;
      predmarg  hse_cat(0) age_cat3(3) dx_time_bin2(0)/adjrr;
run;

proc rlogist data=dedup_thesisx design=srs;
```

```
      class  hse_cat age_cat3 cd4_first_bin1;
      reflev cd4_first_bin1=0 hse_cat=0 age_cat3=3;
      model nevercare= hse_cat age_cat3 cd4_first_bin1;
      predmarg cd4_first_bin1(0) hse_cat(0) age_cat3(3)/adjrr;
run;
*r2 not any higher

*trying some interaction term btwn hse and age;
proc rlogist data=dedup_thesisx design=srs;
      class  hse_cat age_cat3 ;
      reflev  age_cat3=3;
      model nevercare= age_cat3 /*hse_cat  cd4_first_bin1
dx_time_bin2*/ age_cat3*hse_cat;
      predmarg age_cat3(3)*hse_cat(0) /adjrr;
      predmarg age_cat3(3)*hse_cat(1) /adjrr;
run;

/************************FOR LAG MAX******************/
proc sort data=dedup_thesisx out=dedup_thesis2; by nevercare; run;

data dedup_thesis3;
set dedup_thesis2 (keep= nevercare lag_max1 age_cat3 age_cat2 age_cat22
dx_time_bin2 hse_cat race_cat
 cd4_recent_bin curr_gender curr_gender3 trans_categ3 race_cat2);
where nevercare=0;

run;

proc freq data=dedup_thesis3; tables curr_gender*curr_gender2/ list
missing; run;

*bivariate analyses;
proc rlogist data=dedup_thesis3 design=srs;
      class   hse_cat ;
      reflev  hse_cat=0 ;
      model lag_max1= hse_cat;
      predmarg  hse_cat(0)/adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class   age_cat3;
      reflev  age_cat3=3;
      model lag_max1=age_cat3 ;
      predmarg age_cat3(3) /adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class  dx_time_bin2;
      reflev dx_time_bin2=0;
      model lag_max1= dx_time_bin2;
      predmarg  dx_time_bin2(0) /adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class  cd4_recent_bin;
      reflev cd4_recent_bin=0;
      model lag_max1= cd4_recent_bin;
```

```
      predmarg  cd4_recent_bin(0) /adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class  curr_gender3;
      reflev curr_gender3=0;
      model lag_max1= curr_gender3;
      predmarg  curr_gender3(0) /adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class  trans_categ3;
      reflev trans_categ3=0;
      model lag_max1= trans_categ3;
      predmarg  trans_categ3(0) /adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class  race_cat2;
      reflev race_cat2=0;
      model lag_max1= race_cat2;
      predmarg  race_cat2(0) /adjrr;
run;


*full model analysis;
proc rlogist data=dedup_thesis3 design=srs;
      class  hse_cat age_cat3 cd4_recent_bin dx_time_bin2 trans_categ3
curr_gender3 race_cat2;
      reflev cd4_recent_bin=0 hse_cat=0 age_cat3=3 dx_time_bin2=0
trans_categ3=0
            curr_gender3=0 race_cat2=0;
      model lag_max1= hse_cat age_cat3 cd4_recent_bin dx_time_bin2
trans_categ3
            curr_gender3 race_cat2;
      predmarg cd4_recent_bin(0) hse_cat(0) age_cat3(3) dx_time_bin2(0)
curr_gender3(0)
                  trans_categ3(0) race_cat2(0)/adjrr;
run;

*looking at model with selected variables from nevercare;
proc rlogist data=dedup_thesis3 design=srs;
      class  hse_cat age_cat3 cd4_recent_bin dx_time_bin2;
      reflev cd4_recent_bin=0 hse_cat=0 age_cat3=3 dx_time_bin2=0;
      model lag_max1= hse_cat age_cat3 cd4_recent_bin dx_time_bin2;
      predmarg cd4_recent_bin(0) hse_cat(0) age_cat3(3)
dx_time_bin2(0)/adjrr;
run;

*playing around with models;
proc rlogist data=dedup_thesis3 design=srs;
      class   age_cat3 cd4_recent_bin dx_time_bin2;
      reflev cd4_recent_bin=0  age_cat3=3 dx_time_bin2=0;
      model lag_max1=  age_cat3 cd4_recent_bin dx_time_bin2;
      predmarg cd4_recent_bin(0)  age_cat3(3) dx_time_bin2(0)/adjrr;
run;
```

```
proc rlogist data=dedup_thesis3 design=srs;
      class   age_cat3 dx_time_bin2;
      reflev  age_cat3=3 dx_time_bin2=0;
      model lag_max1=  age_cat3  dx_time_bin2;
      predmarg   age_cat3(3) dx_time_bin2(0)/adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class  hse_cat  dx_time_bin2;
      reflev  hse_cat=0  dx_time_bin2=0;
      model lag_max1= hse_cat  dx_time_bin2;
      predmarg hse_cat(0)dx_time_bin2(0)/adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class   cd4_recent_bin dx_time_bin2;
      reflev cd4_recent_bin=0 dx_time_bin2=0;
      model lag_max1= cd4_recent_bin dx_time_bin2;
      predmarg cd4_recent_bin(0) dx_time_bin2(0)/adjrr;
run;

proc rlogist data=dedup_thesis3 design=srs;
      class   age_cat3 cd4_recent_bin;
      reflev cd4_recent_bin=0  age_cat3=3;
      model lag_max1=  age_cat3 cd4_recent_bin;
      predmarg cd4_recent_bin(0)  age_cat3(3)/adjrr;
run;
```

```
/*************************************************************************

Author: Katrece Outlaw
Purpose: SJMC Matches logistic regression analysis for OR
Date:   February 2015
Location: H:\Thesis
Last Modified:

*************************************************************************/

dm log 'clear';
dm output 'clear';

options nocenter nodate nonumber nofmterr mprint mlogic symbolgen;

libname stj 'H:\Thesis';

/**************Logistic Regression Bivariate Analyses**************/


proc logistic data=dedup_thesis;
class race_cat (ref='W')/ param=ref;
model nevercare (event='1') = race_cat;
run;

proc logistic data=dedup_thesis;
class age_cat3 (ref='3')/ param=ref;
model nevercare (event='1') = age_cat3;
run;

proc logistic data=dedup_thesis;
class curr_gender (ref='M')/ param=ref;
model nevercare (event='1') = curr_gender;
run;

proc logistic data=dedup_thesis;
class curr_gender2 (ref='M')/ param=ref;
model nevercare (event='1') = curr_gender2;
run;

proc logistic data=dedup_thesis;
model nevercare (event='1') = sqcd41;
run;

proc logistic data=dedup_thesis;
model nevercare (event='1') = cd4_first_bin;
run;

proc logistic data=dedup_thesis;
model nevercare (event='1') = hse_cat;
run;

proc logistic data=dedup_thesis;
class trans_categ2 (ref='other')/ param=ref;
model nevercare (event='1') = trans_categ2;
run;
```

```
proc logistic data=dedup_thesis;
model nevercare (event='1') = dx_time_bin2;
run;

*look at interaction term;
proc logistic data=dedup_thesis;
class age_cat3 (ref='3')/ param=ref;
model nevercare (event='1') = hse_cat age_cat3 hse_cat*age_cat3;
run;

/*******************Multivariable Logistic Regression**************/

proc logistic data=dedup_thesis;
model nevercare (event='1')= hse_cat dx_time_bin;
run;

*looking at variables all binary;
proc logistic data=dedup_thesis;
model nevercare (event='1')= hse_cat age_cat22 cd4_first_bin1
dx_time_bin2/
selection= /*backward*/ none
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;

*looking at variables all categories;
proc logistic data=dedup_thesis;
class age_cat (ref="4") cd4_first_cat (ref="6") dx_time_cat (ref="3")/
param=ref;
model nevercare (event='1')= hse_cat age_cat cd4_first_cat dx_time_cat/
selection=none
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;

*looking at variables all continuous;
proc logistic data=dedup_thesis;
class hse_cat (param=effect);
model nevercare (event='1')= hse_cat cur_age1 cd4_first_value dx_time/
selection=backward
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;

*looking at variables mixed cat and continuous;
proc logistic data=dedup_thesis;
class hse_cat (param=effect) age_cat (ref="4") cd4_first_cat (ref="6")/
param=ref;
```

```
model nevercare (event='1')= hse_cat age_cat cd4_first_cat dx_time/
selection=backward
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;

*looking at variables mixed binary category and continuous;
proc logistic data=dedup_thesis;
class cd4_first_cat (ref="6")/ param=ref;
model nevercare (event='1')= hse_cat age_cat22 cd4_first_cat dx_time/
selection= backward
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;

*full models;
proc logistic data=dedup_thesis;
class race_cat (ref="W") curr_gender2 (ref="M") trans_categ2
(ref="other")  age_cat3 (ref='3')/ param=ref;
model nevercare (event='1')= race_cat curr_gender2 hse_cat age_cat3
      trans_categ2 cd4_first_bin dx_time_bin2 hse_cat*age_cat3/
selection= stepwise
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;

*full, no selection;
proc logistic data=dedup_thesis;
class race_cat (ref="W") curr_gender2 (ref="M") trans_categ2
(ref="other") age_cat3 (ref='3')/ param=ref;
model nevercare (event='1')= race_cat curr_gender2 hse_cat age_cat3
      trans_categ2 cd4_first_bin /*sqcd41*/ dx_time_bin2
hse_cat*age_cat3/
selection= none
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;

*playing around with variables to include in the model;
proc logistic data=dedup_thesis;
class age_cat3 (ref='3');
model nevercare (event='1')= age_cat3 hse_cat cd4_first_bin
dx_time_bin2 hse_cat*age_cat3;
run;
```

```
proc logistic data=dedup_thesis;
class  age_cat3 (ref='3')/ param=ref;
model nevercare (event='1')=  hse_cat age_cat3 dx_time_bin2
cd4_first_bin/
selection= none
            SLE=0.05
            SLS=0.05
            INCLUDE=0
            LINK=LOGIT
            ALPHA=0.05;
run;
```

Appendix B: SAS Code Variable Descriptions

Variable Name	Description	Coding	Format	Comments
age_cat	current age in 5 categories	1= 18-24 2= 25-34 3=35-44 4=45-54 5= 55+	Num	created from cur_age1
age_cat3	current age in 3 categories		Num	created from cur_age1
aids_age_yrs	Age at AIDS diagnosis (years)		Num	original from ehars
aids_cdc	CDC case definition for AIDS	Y=yes N= no	YN	original from ehars
aids_insurance	Primary reimbursement for medical treatment (AIDS)		Char	original from ehars
birth_sex	Sex at birth		Char	original from ehars
cd4_recent_cnt_dt	Most recent CD4 count test result date		Char	original from ehars
cd4_recent_dt	Most recent CD4 count test result date		Char	created from above variable. Added 15 as day to dates missing a day value
_cd4_recent_dt	Most recent CD4 count test result date		Num/ YYMMDD10.	created from above variable. Converted to SAS numeric date.
cd4_recent_cnt_value	Most recent CD4 count test result value		Char	original from ehars
cd4_recent_cnt	Most recent CD4 count test result value		Num	created from above variable. Converted to numeric cnt
cd4_recent_bin	Binary variable for most recentlyy recorded CD4 lab count	0= >200 cells/mm3; 1= ≤ 200 cells/mm3	Num	created from above variable.
cd4_vl_first_hiv_dt	First CD4 or viral load test result date after HIV diagnosis		Char	original from ehars
cd4_vl_first_dt	First CD4 or viral load test result date after HIV diagnosis		Char	created from above variable. Added 15 as day to dates missing a day value
_cd4_vl_first_dt	First CD4 or viral load test result date after HIV diagnosis		Num/ YYMMDD10.	created from above variable. Converted to SAS numeric date.
cd4_first_hiv_value	First CD4 test result value after HIV diagnosis		Char	original from ehars
cd4_first_value	First CD4 test result value after HIV diagnosis		Num	created from above variable. Converted to numeric cnt
cd4_first_bin	Binary variable for first recorded CD4 lab count	0= >200 cells/mm3; 1= ≤ 200 cells/mm3	Num	created from cd4_first_value
cur_age	Current age (calculated using the date of analysis data set creation)		Char	original from ehars
cur_age1	current age		Num	created from above variable. Converted to numeric value.
current_gender	current gender	M = male, F=female, MF= transgender, U=unknown	Char	original from ehars
curr_gender2	same	M = male, F=female, MF= transgender,	Char	created from current_gender making U = missing
curr_gender3	Numeric category variable for gender	0= male, 1= female, 2= MF transgender	Num	created from curr_gender2

Variable Name	Description	Coding	Format	Comments
disease_categ_dx	Disease progression (diagnosis date)		Char	original from ehars
dob	date of birth		Char	original from ehars
dx_time	diagnosis time, time from hiv diagnosis date to 4/21/2014		Num	calculated from hiv_dx_date
dx_time_bin2	Binary variable for diagnosis time	0= >10 years; 1= ≤ years		
ethnicity1	ethnicity	E1= hisp, E2= not hisp UNK=unknown	Char	original from ehars
first_name	First Name		Char	original from ehars
hiv_aids_age_yrs	Age at HIV disease diagnosis (years)		Char	original from ehars
hiv_aids_age	Age at HIV disease diagnosis (years)		Num	created from above variable. Converted to numeric value.
Homelessstatus	Housing status of pt encounters from STJ		Char	original from STJ Encounters data set
hsing_status	Housing status of pt encounters from STJ	1=not homeless, 2=shelter/hotel/boarding house 3= street/transitional, 4=w/friends or relatives 5=halfway hse/substance abuse facility, 6=jail, 7=psych/hospital, 8=domestic violence shelter, 9=unknown	Num	created from homelesstatus
hse_cat	Binary variable for housing status	0 = not homeless; 1= homeless (2-9 from hsing_status variable)	Num	created from hsing_status
ID	unique patient identifier		Num	created from document_uid by adding random number
lag_day	# days between subsequent lab tests		Num	created using lag function; only seen on data set containing duplicates
lag2_day	# days between subsequent 2 lab tests		Num	created using lag function; only seen on data set containing duplicates
lag_max	max # days between subsequent lab tests		Num	created by finding max value from all lag values per each doc_uid. Only on de-dup data set
lag_max1	indicator variable for having a gap in care	1= lag of >180 days between lab dates	Num	
lag_mo	# months between susbsequent lab tests		Num	created using lag function; only seen on data set containing duplicates
last_enc_yr	Year of last encounter date at SJMC		Num	taken from original STJ encounter data
last_name	Last Name		Char	original from ehars
most_recent	Most recent test date (CD4 or VL)		Num/ YYMMDD10.	Created from _cd4_recent_dt and _vl_recent_dt

Variable Name	Description	Coding	Format	Comments
nevercare	Indicator variable for never received care	1= most recent test date within 3 months of first test date or no recorded test lab date.	Num	created using most_recent and _cd4_vl_first_dt
race	duh	1= Hispanic, any 2= Native American/AI, 3=Asian, 4=AA/Black, 5=Hawaiian/PI, 6=White, 7=Legacy Asian/ PI, 8=multi-racial, 9=unknown	Num	original from ehars
race_cat	category variable for race	B= Black, H= Hispanic, any, W= White, O=Other	Char	created from race variable
race_cat2	numeric category variable for race	0= White, 1= Black, 2= Hispanic, 3= Other	Num	created from race_cat variable
result	result of test taken at sample_dt		Char	original from ehars
sample_dt	date of test sample		Char	original from ehars lab doc data
sample_dt2	date of test sample		Char	created from above variable. Added 15 as day to dates missing a day value
_sample_dt2	date of test sample		Num/ YYMMDD10.	created from above variable. Converted to SAS numeric date.
T1days	# days from 4/21/2014 until most recent test date		Num	created by substracting most_recent from 4/21/2014
T1exd	# days from out of care window until most recent test date		Num	created by substracting 455 days from T1days
T1exy	# years from out of care window until most recent test date		Num	created by substracting 455 days from T1yrs
T1yrs	# years from 4/21/2014 until most recent test date		Num	
trans_categ	transmission category	1= (MSM) 2= (IDU) 3 = (MSM+IDU) 4= Adult receipt of clotting factor blood product 5= Heterosexual contact with someone who had HIV infection or had an MSM or IDU risk factor for it 6= Adult receipt of transfusion of blood or blood components, transplant of organ or tissue, or artificial insemination 7= Perinatal (mother-to-child) exposure 8= Adult with other risk factor such as occupational exposure 9= Adult with no identified risk factor (NIR) (³12 months after diagnosis) 10= Adult with no reported risk factor (NRR) (<12 months after diagnosis)	Char	orignal from ehars
trans_categ2	Category variable for transmission category	other= 2, 4-9 from above; MSM =1 and 3 from above	Char	created from trans_categ
trans_categ3	Binary variable for transmission category	0= other; 1=MSM	Num	created from trans_categ2
vl_recent_cnt	Most recent viral load test result value (copies/ml)		Num	numeric count created from value variable
vl_recent_value	Most recent viral load test result value (copies/ml)		Char	original from ehars

Appendix C: Full Multivariable Logistic Regression Models

Table 4. Multivariable Logistic Regression Analysis of Never in HIV Care* with Demographic and Clincal Variables N=395

Variable	PR (95% CI)¥	p - value
Race		
White	Referent	
Black	0.71 (0.45, 1.12)	0.1651
Hispanic	0.61 (0.28, 1.33)	0.1840
Other	0.30 (0.09, 0.98)	0.0255
Gender		
Male	Referent	
Female	0.42 (0.19, 0.94)	0.0275
Transgender	1.77 (1.11, 2.81)	–
Housing		
Not Homeless	Referent	
Homeless	1.75 (1.23, 2.48)	0.0225
Age		
> 35	Referent	
26-35	1.30 (0.84, 2.02)	0.2349
18-25	2.66 (1.55, 4.58)	0.0064
Risk		
Other	Referent	
MSM^	0.67 (0.42, 1.06)	0.1085
Diagnosis Time°		
> 10 yrs	Referent	
≤ 10 yrs	2.88 (1.72, 4.82)	<0.0001

* Never engaged in care defined as no recorded CD4+ T lymphocyte or HIV viral load lab data or most recent lab data within 3 months of HIV diagnosis date.

¥ Prevalence Ratio with associated 95% confidence interval

^ MSM, men who have sex with men

° Defined as time from HIV diagnosis date to April 21, 2014

Table 5. Multivariable Logistic Regression Analysis of Gap in HIV Care with Demographic and Clincal Variables, N= 286**

Variable	PR (95% CI)¥	p - value
Race		
White	Referent	
Black	1.29 (0.94, 1.78)	0.0484
Hispanic	1.47 (1.03, 2.08)	0.0458
Other	1.26 (0.80, 1.98)	0.3453
Gender		
Male	Referent	
Female	0.96 (0.65, 1.40)	0.8065
Transgender	0.64 (0.44, 0.94)	–
Housing		
Not Homeless	Referent	
Homeless	1.00 (0.86, 1.17)	0.9913
Age		
> 35	Referent	
26-35	0.86 (0.69, 1.07)	0.1145
18-25	1.01 (0.63, 1.60)	0.9748
Risk		
Other	Referent	
MSM^	0.92 (0.69, 1.23)	0.5937
Diagnosis Time°		
> 10 yrs	Referent	
≤ 10 yrs	0.89 (0.76, 1.05)	0.1372
Recent CD4€		
> 200	Referent	
≤ 200	1.02 (0.86, 1.22)	0.7916

** Defined as >180 days between any 2 lab dates

¥ Prevalence Ratio with associated 95% confidence interval

^ MSM, men who have sex with men

° Defined as time from HIV diagnosis date to April 21, 2014

€ Most recent CD4 T lymphocyte lab count recorded