Estimating Prevalent Diagnoses and Rates of New Diagnoses of HIV at the State Level by Age Group Among Men who Have Sex With Men in the United States

Jeb Jones,1 Jeremy A. Grey,2 David W. Purcell,2 Kyle T. Bernstein,2 Patrick S. Sullivan,1 and Eli S. Rosenberg3

1Department of Epidemiology, Emory University, and 2National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia; and 3Department of Epidemiology and Biostatistics, University at Albany, Albany, New York

Background. Men who have sex with men (MSM) in the United States experience a disproportionate rate of diagnosis of HIV. Surveillance data demonstrate age-based disparities among MSM, with higher rates of diagnosis among MSM age ≤34 years nationally. Population size estimates within age group at the state level have not been available to determine rates for each state. We estimated the size of the MSM population in 5 age groups in each state and estimated the rate of prevalent HIV diagnoses in 2013 and new HIV diagnoses in 2014.

Methods. We used data from the General Social Survey, American Community Survey, and previously published estimates from the National Health and Nutrition Examination Survey to estimate the population of MSM in 5 age groups at the state level. We combined these estimates with surveillance data to estimate age-stratified rates of prevalent diagnoses in 2013 and new diagnoses in 2014 in each state. We estimated standardized prevalence and diagnosis ratios comparing the Northeast, South, and West regions with the Midwest.

Results. Rates of prevalent diagnoses increased with increasing age, and rates of new diagnoses were highest among younger age groups. In the United States, the new diagnosis rate among those age 18–24 years in 2014 was 1.4 per 100 MSM without a diagnosis. The highest diagnosis rates were observed among men age ≤34 years in the South.

Conclusions. Age-stratified estimates of HIV prevalence and new diagnosis rates at the state level can inform public health prevention strategies and resource allocation.

Keywords. demography; HIV; men who have sex with men.

Men who have sex with men (MSM) in the United States experience a disproportionate number and highly disproportionate rate of new HIV diagnoses annually compared with other risk groups [1–3]. Among MSM, the distribution of new diagnoses is not consistent across subgroups of MSM; disparities in incidence have been noted with respect to age and race [4]. Nationally, rates of new HIV diagnoses are highest among MSM younger than age 35 years and black and Hispanic MSM [4, 5]. There is also significant geographic variability in rates of new diagnoses of HIV and of HIV prevalence [5]. For example, the South is the only region with state-level prevalent HIV diagnoses >15% among MSM and rates of new diagnoses >1.0/100 MSM per year [6].

A consistent challenge in HIV surveillance has been the lack of information about the sizes of the MSM population and subpopulations across the United States. This is due to multiple factors, including that these populations are not enumerated in the US Census. Yet these estimates are necessary to compute denominators for rates, which contextualize the HIV epidemic. Estimates of the number of MSM have been produced at the national [4, 7], state [8–10], and county levels [10–12], using new statistical approaches applied to representative data sources. These estimates make it possible to calculate rates of HIV diagnosis among MSM in specific geographic regions and racial subgroups to better understand the HIV epidemic and to guide public health decision-making in terms of targeting HIV prevention resources where they will have the biggest impact.

Age-stratified estimates of MSM in the United States have allowed for calculation of rates of new HIV diagnosis by age at the national level [13]. Both absolute counts and diagnosis rates show that the highest number of diagnoses occur in MSM age ≤34 years in 2015 [5]. However, age-stratified estimates have been unavailable for smaller geographic areas. Estimating the population size of each age group at the state level will help better characterize the HIV epidemic in the United States and in smaller geographic areas in terms of prevalence and new diagnoses.

We extended our previously described approach for estimating population sizes of MSM [10] to develop age-stratified estimates of the MSM population in each US state and the
District of Columbia (hereafter: states) in 2013 and 2014. These estimates were then combined with HIV surveillance data to determine state-level prevalence of HIV diagnosis in 2013 and new diagnosis rate in 2014 among MSM in each age group at the state level.

METHODS

Denominator Data Sources and Methods

We extended our previous method of estimating the number of total MSM in US counties, described in detail elsewhere [10, 14], to further estimate the number of MSM in 2013 and 2014 in 5 age categories: 18–24, 25–34, 35–44, 45–54, and 55 years and older. We began by using the original method to estimate the total number of MSM in each county. Briefly, the number of MSM was determined by the county’s US Census region, National Center for Health Statistics urban–rural classification [15], density of male–male households, and number of male residents age 18 years or older. We began by multiplying regional and urbanicity-specific estimates of the percentage of MSM among adult men [16], based on reported sexual behavior, by the number of adult men in each county. We then calculated the county-level percentage of households that had a male head of household and a male domestic partner, using data from the American Community Survey (ACS) 5-year Summary Files (2009–2013 for 2013 estimates, 2010–2014 for 2014 estimates) [17, 18], to obtain the county’s density of male–male households. The ratio of each county’s density of male–male households to the overall density among all counties at the same level of urbanicity within the same region was used to adjust the percentage of MSM among adult men, allocating a higher percentage of MSM to counties with relatively higher concentrations of male–male households. This adjusted percentage was multiplied by the number of adult men. Finally, all county MSM populations were scaled to equal an aggregated total of MSM in the United States that matched the estimated national percentage of men who had sex with another man in the past 5 years [4].

Following the estimation of county-level MSM population sizes, we applied an estimated age distribution of MSM to obtain age-specific estimates. Estimating age distributions of MSM involved a 3-step process. First, we used pooled data from the 1972 through 2014 waves of the General Social Survey (GSS) [19] to estimate same-sex behavior in each of our target age groups, stratified by 4 levels of urbanicity. The percentage of men reporting sex with a man in the past 5 years within each stratum of age group and urbanicity was calculated using sample weights that accounted for changes to sampling methods in the GSS over time, including periodic oversampling of black populations and varying sample designs (ie, “block quota” or “full probability” designs) [20]. The percentage of men with a male sex partner in the past 5 years by age group and urbanicity are presented in Supplementary Table 1. We then multiplied the weighted percentages by age-specific male population in each county from the ACS [17, 18], according to county urbanicity level, yielding the age distribution of total MSM who had had sex with a man in the previous year (Equation 1).

\[
\text{Percentage of } \text{MSM}_{\text{age group}} \times \text{urbanicity} \times \text{Male population of county}_{\text{age group}} \times \text{age distribution of } \text{MSM}
\]

We then calculated the percentage of MSM in the county who fell into each age category. Finally, we multiplied the percentage of MSM in each age group in the county by the estimated number of MSM in the county obtained using the original method. County-level estimates were then summed to obtain state-level estimates. Sample calculations to determine the age-specific MSM population for an example county are provided in Supplementary Table 2.

As with our previous work [14], we estimated variability around each estimate by conducting 100 000 repetitions of a bootstrapping method that used input values randomly drawn from error distributions around data sources. For the age-specific MSM estimates, this approach accounted for variability in the ACS (ie, male population by age and housing data at the county level) [17, 18], published estimates of the percentage of MSM among adult men [16–18], and the percentage of MSM among men in each age category from the GSS [19]. We then used medians from simulated distributions as denominators, with the 2.5th and 97.5th percentiles as 95% confidence interval bounds, for rate calculations.

HIV Surveillance Data Sources

Age-stratified counts of MSM living with a diagnosis in 2013 and with a new diagnosis in 2014 in each state were obtained from AIDSVu.org [21], which provides HIV surveillance data collected by state and local health departments that have been compiled and processed by the Centers for Disease Control and Prevention (CDC). Age-stratified estimates of prevalent HIV diagnoses among MSM (2013) and new diagnoses (2014) were not available for New Hampshire. HIV surveillance statistics for states with fewer than 12 cases in an age stratum were suppressed but still contributed to overall state, division, and region estimates.

Analysis

Age-stratified state-level prevalence of living with an HIV diagnosis was determined by dividing prevalent HIV diagnoses for MSM in 2013 by the estimated population of MSM in each age category for that year.

Rates of new diagnosis for 2014 were calculated among MSM not living with an HIV diagnosis, who were thus at risk of receiving an HIV diagnosis in 2014. This provides an epidemiological measure analogous to incidence, substituting new diagnoses for new infections [6]. MSM at risk in 2014 were obtained by subtracting age-specific prevalent diagnoses among MSM in 2013 from the population of MSM in each age stratum in each state in 2014. New diagnoses for 2014 were then divided by the estimated population of MSM at risk.
Variability measures are not available for diagnoses; therefore, we used the 95% confidence intervals from the denominators to estimate 95% confidence intervals for all HIV indicators. Upper and lower bounds for the HIV indicators were obtained by dividing the surveillance data by the lower and upper bounds of the population estimates, respectively.

Standardized prevalence ratios (SPRs) and standardized diagnosis ratios (SDRs) were calculated to compare the Northeast, South, and West regions with the Midwest, adjusting for age. The Midwest region was selected as the referent because it tends to have lower prevalence and new diagnoses compared with other regions. To calculate SPRs for each age group, the prevalence rate for the Midwest region was applied to the estimated 2013 population of the comparison region, yielding prevalence rates for the comparison region under the null hypothesis of no regional difference. The ratio of observed-to-expected prevalence was then calculated to estimate the SPR. Similarly, to determine the SDR for new diagnoses, the new diagnosis rate for the Midwest region was applied to the estimated 2014 population of at-risk MSM in the comparison region. Confidence intervals were determined via the Fisher exact method.

Analyses were conducted using R (v3.3.3), Microsoft Excel, and Openepi.com.

RESULTS

MSM Population by Age and State

Figure 1 displays the distribution of MSM in each state in 2013 across the 5 different age groups that were modeled: 18–24 years, 25–34 years, 35–44 years, 45–54 years, and 55 years and older.

The size of the MSM population generally tracked the overall population sizes of the states, with the states with the largest populations consistently ranked highest in population of MSM across the age groups.

2013 Prevalence of Living With an HIV Diagnosis

Age-stratified estimates of prevalent HIV diagnoses for each Census region, division, and state are presented in Figure 2 and Supplementary Table 3. Across the strata of age groups and states, prevalence ranged from as low as 0.2% of MSM in the 18–24-year age group to 36.4% in the 25–34-year age group. Prevalence tended to increase within a state through the 45–54-year age group, often declining among those age 55 years and older. The South was the only region with states that had HIV prevalence >10% among 18–24-year-olds, with 3 states in the region having prevalence >10% in this age group. Montana and Wyoming were the only 2 states with prevalence below 5% among MSM in all age groups. Four states had at least 1 age-specific HIV prevalence >30% among MSM: New York (45–54 years, 55 years and older), Mississippi (25–34 years, 45–54 years), South Carolina (45–54 years), and Washington, DC (45–54 years). SPRs (Table 1) indicate that, compared with the Midwest, the Northeast and the South had higher than expected prevalence in all age groups. The South had the highest SPR in the 18–24-year age group (SPR, 1.41; 95% confidence interval [CI], 1.29–1.44), and the West had the highest SPR in the remaining age groups, with the highest among those age 55 years and older (SPR, 2.20; 95% CI, 2.17–2.23). The West had lower than expected prevalence in younger age groups and higher than expected prevalence among older age groups.
Rates of new diagnoses in 2014 are presented among all MSM (Figure 3) and MSM without a diagnosis in 2013 (Figure 4). Full data are presented in Supplementary Table 4. New diagnosis rates were generally highest among the 18–24-year age range and were lower in subsequently higher age brackets. Rates ranged from a low of 0.0 (18–24 years, Wyoming; 45–54 years, South Dakota and Montana; 55 years and older, Montana, North Dakota, and Wyoming) to a maximum of 6.5 per 100 MSM age 25–34 years at risk in Mississippi. Louisiana, Mississippi, and South Carolina each had age-specific diagnosis rates >5.0 per 100 MSM at risk for at least 1 age group. Seven states had diagnosis rates <0.5 per 100 MSM at risk for all 5 age categories: Idaho, Montana, Maine, Oregon, South Dakota, Utah, and Washington. Standardized diagnosis ratios indicated higher than expected diagnosis rates in the Northeast and South in all age groups and in the West in older age groups compared with the Midwest. The highest SDRs among men age 18–24 years (SDR, 1.60; 95% CI, 1.55–1.65) and 25–34 years (SDR, 1.75; 95% CI, 1.71–1.80) occurred in the South; the Northeast had the highest SDRs in the older age groups. Overall age-adjusted SDRs in the Northeast and South indicated 70% (SDR, 1.70; 95% CI, 1.65–1.75) and 74% (SDR, 1.74; 95% CI, 1.71–1.77) more diagnoses than expected compared with the Midwest, respectively.

Diagnosis rates among MSM at risk are plotted in Figure 4. This graph depicts the diagnosis rate of HIV among MSM at risk based on the size of the MSM population in each state, grouped by region. The size of each data point represents the size of the MSM population at risk in that age group. This illustrates that although rates of new diagnoses were highest in the South, these states had relatively small populations of MSM at risk.

### DISCUSSION
This is the first study to estimate age-specific HIV prevalence and diagnosis rates at the state level in the United States. By...
subtracting prevalent cases from the 2014 MSM population, we were also able to estimate the diagnosis rate among MSM at risk in each state instead of among the full MSM population, as has been done in previous studies [4].

The highest rates of new diagnoses were observed among MSM age 18–24 years, a finding that was consistent across all regions. This reflects ongoing trends in the United States that have been noted at the national level [4, 13] indicating that this...
age group experiences disproportionate rates of new diagnoses, which is at least partly explained by having a higher proportion of uninfected persons than older-age cohorts. The largest SDRs among men younger than age 34 years were observed in the South, followed by the Northeast. The states in the South with the highest rates of HIV also had relatively small MSM populations, indicating that these states have smaller numbers of at-risk men than some other states but that these men are at considerable HIV risk.

Numerous HIV prevention modalities are available, including condoms [22], pre-exposure prophylaxis (PrEP) for HIV-negative persons [23, 24], and HIV treatment with antiretroviral therapy for persons living with HIV, which reduces the risk of transmission to negligible levels if viral suppression is achieved and maintained [25]. The results of the current analysis suggest that programs promoting viral suppression for HIV-positive persons and other appropriate prevention methods such as PrEP and condoms for HIV-negative persons should be expanded in the South and Northeast and should be targeted to MSM at younger ages. Younger men living with HIV have worse outcomes across the HIV continuum of care [26] and may not have well-established prevention repertoires to protect themselves and their partners. Messaging around HIV prevention should target the appropriate audience, and different messaging strategies are needed for different audiences and age groups [27]. Knowledge of which groups are most affected in certain areas will also help to appropriately target messaging campaigns.

HIV prevalence continues to grow over the age groups as HIV diagnoses accumulate over time, with the highest prevalence observed among MSM older than age 45 years. This population of men living with HIV will have different needs [28] compared with younger men. Continued efforts to understand how to serve a population aging with HIV both in terms of their needs related to maintaining viral suppression and addressing comorbidities associated with HIV and aging will be necessary.

We observed higher diagnosis rates in the South and Northeast compared with the Midwest in each age group. Continued efforts to understand the causes of these geographic disparities are needed. Among MSM age 18–24 and 25–35 years—the age groups with the most HIV diagnoses—the highest SDRs were observed in the South, followed by the Northeast, indicating that high diagnosis rates among young MSM are predominantly clustered in the Southern and Northeastern United States. Elevated prevalence and diagnosis rates have previously been reported among MSM in the South [29]. This suggests that HIV prevention programs in the South and Northeast should focus on young MSM. The United States has a national prevention goal of reducing HIV infections among disproportionately affected populations, including MSM and people in the Southern United States [30]. MSM in the Deep South in particular continue to bear a disproportionate burden of new diagnoses and prevalent cases—a trend that is most notable among the youngest age groups [31]. The CDC allocates funding for HIV prevention to target the populations most affected, including increasing funding to health departments and community-based organizations for HIV prevention programs in the South [32, 33]. Continued targeting in these states is needed to reduce diagnosis rates and increase linkage to care.

Hess et al. [13] have developed a method for assessing the lifetime risk of HIV in the United States overall and among different subpopulations. Their estimates rely on the strong and previously untested assumption that different age cohorts will have the same pattern of age-specific HIV diagnosis rates over time and that national rates of age-specific MSM diagnoses are representative across the country. Our results indicate that these diagnosis rates differ substantially across states within a given year. The methods presented in this paper provide a means of tracking changes in diagnosis rate within an age group over time to inform improved estimates of lifetime risk. Additionally, the wide variation observed across states suggests that, overall, national estimates of lifetime risk would not be informative with respect to any specific MSM population in the United States. Rather, such estimates might need to be calculated for smaller geographical areas.

State-level estimates of HIV prevalence within each age group of MSM inform public health resource allocation for HIV prevention and treatment programs. Prevalence estimates provide an indication of the expected health care burden associated with linking and maintaining all persons living with HIV (PLWH) in care. The complement of prevalence estimates provides an estimate of the size of the population of MSM at risk of HIV acquisition, which is useful for planning HIV prevention programs.

Each stage of the HIV prevention [34] and care [35] continua requires a different type of intervention, from condom distribution, HIV testing, and PrEP prescription for those at risk to initiation and retention in medical care to viral suppression for those living with HIV [36, 37]. Understanding what these different age-related views of the HIV epidemic mean at the state level will allow more efficient allocation of HIV prevention and treatment resources. The CDC publishes the “State HIV Prevention Progress Report” summarizing progress within each state toward national goals of HIV prevention, increasing access to HIV care, improving HIV care outcomes, and reducing HIV-related health disparities [38]. The ability to track HIV prevalence and diagnosis rates at the state level by age group for MSM complements these existing goals to provide further evidence of successes and challenges in HIV prevention and treatment programs.

Despite the novelty of our methods, it is important to address the limitations of these analyses. First, although race is associated with HIV prevalence and diagnosis rate, with black and Hispanic MSM experiencing higher rates of diagnosis
compared with white MSM [4], we did not account for race of MSM. Race-stratified estimates of prevalent and new HIV diagnoses have been generated [39]; however, 3-way stratification (risk group × race × age) of HIV surveillance data is not publicly available at the state level, and stratification at this level would result in unstable estimates in states with small population sizes of some races. Second, we used multiple sources of data to generate estimates of the MSM population; each of these data sources is subject to potential error and misclassification. Some of these limitations would be overcome with more direct measures of MSM sexual behavior and identity in national surveys. For example, the estimates of MSM population sizes in each age group are based on estimates from the GSS of the percentage of MSM in each age group, based on urbanicity. The proportion of MSM in the GSS is based on a 5-year recall of same-sex sexual behavior among male respondents. If men in younger age groups have sex more frequently, then this would bias the GSS estimates, resulting in overestimates of the MSM population in younger age groups and underestimates among older age groups. Given that we observed high rates of new diagnoses among younger MSM and high prevalence of living with a diagnosis among older MSM, such biases appear to be minimized. Third, ACS data are from a weighted sample and are thus less comprehensive than data from the decennial census. Additionally, these data rely on male–male households as a proxy measure of MSM. Underreporting of same-sex male households might result if such households in certain areas are less likely to respond. Third, differences in the legal status of same-sex relationships during the reporting period (ie, 2013–2014) might affect accurate ascertainment of the number of same-sex male households. Fourth, some of the data sources used to estimate the MSM population rely on self-report of sexual behavior. Nondisclosure of same-sex sexual behavior would result in underestimates of the MSM population and overestimates of new diagnoses and prevalent diagnoses. Fifth, the surveillance data used to estimate prevalent diagnoses and diagnosis rate are collected by state health departments. Differences in the completeness of data reporting and programs that increase recognition of people living with HIV (eg, HIV testing promotion) across jurisdictions might lead to differential ascertainment of HIV cases. However, HIV surveillance systems are periodically evaluated for completeness of reporting, and national surveillance oversight sets minimum standards for completeness of reporting. Finally, diagnosis rates are a lagging indicator. Diagnosis might occur years after infection; thus, high diagnosis rates could be an indication of high incidence or of effective testing efforts that increase detection of existing cases, or both.

Overall, we observed a wide range of variability across states and age groups in terms of HIV diagnosis rates among MSM at risk and MSM living with HIV. However, across states and regions, we found the highest rates of new HIV diagnosis among MSM age 18–24 years. Had we only examined absolute numbers of new diagnoses (not rates), 25–34-year-old MSM would have the largest amount of disease. These 2 different epidemiologic measures represent different views of the HIV epidemic and may identify different needs in terms of prevention and treatment resources. The age-specific prevalence and diagnosis rates presented in this study represent a step forward in describing and understanding the HIV epidemic among MSM in the United States and will allow policy-makers, researchers, and public health program administrators to plan and implement more responsive HIV prevention and treatment programs.

Supplementary Data
Supplementary materials are available at Open Forum Infectious Diseases online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Acknowledgments
We thank members of the scientific and public health advisory groups of the Coalition for Applied Modeling for Prevention project for their input on this study, and specifically those members who reviewed a previous version of this manuscript: Thomas Bertrand, Mary Ann Chiasson, David Dowdy, Gregory Felzien, David Holtgrave, and Jane Kelly.

Financial support. This work was supported by the Centers for Disease Control and Prevention National Center for HIV, Viral Hepatitis, STDs, and TB Prevention Epidemic and Economic Modeling Agreement (NEEMA; U38 PS004646-01).

Disclaimer. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Potential conflicts of interest. The authors have no conflicts to disclose.

References
1. Centers for Disease Control and Prevention. HIV surveillance report, 2013. Vol. 25. 2015. https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2013-vol-25.pdf. Accessed 15 July 2017.
2. Johnson AS, Hall HI, Hu X, et al. Trends in diagnoses of HIV infection in the United States, 2002-2011. JAMA 2014; 312:432–4.
3. Hall HI, Song R, Rhodes P, et al; HIV Incidence Surveillance Group. Estimation of HIV incidence in the United States. JAMA 2008; 300:520–9.
4. Purcell DW, Johnson CH, Lansky A, et al. Estimating the population size of men who have sex with men in the United States to obtain HIV and syphilis rates. Open AIDS J 2012; 6:98–107.
5. Centers for Disease Control and Prevention. HIV surveillance report, 2015. Vol. 27. 2016. https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2015-vol-27.pdf. Accessed 15 August 2017.
6. Rosenberg ES, Grey JA, Sanchez TH, Sullivan PS. Rates of prevalent HIV infection, prevalent diagnoses, and new diagnoses among men who have sex with men in US states, metropolitan statistical areas, and counties, 2012-2013. JIMIR Public Health Surveill 2016; 2:e22.
7. Oster AM, Sternberg M, Lansky A, et al. Population size estimates for men who have sex with men in the United States, metropolitan statistical areas, and counties, 2012-2013. JIMIR Public Health Surveill 2016; 2:e22.
8. Lieb S, Fallon SJ, Friedman SR, et al. Statewide estimation of racial/ethnic populations of men who have sex with men in the U.S. Public Health Rep 2011; 126:60–72.
9. Lieb S, Thompson DR, Mirsa S, et al. Southern AIDS Coalition MSM Project Team. Estimating populations of men who have sex with men in the southern United States. J Urban Health 2009; 86:887–901.
10. Grey JA, Bernstein KT, Sullivan PS, et al. Estimating the population sizes of men who have sex with men in US states and counties using data from the American Community Survey. JIMIR Public Health Surveill 2016; 2:e14.
11. Campagna J, Poe J, Robbins A, Rowlinson E. Estimating the number of men who have sex with men by race/ethnicity at the county level in Texas. J Urban Health 2015; 92:168–81.
