

Published in final edited form as:

*J Gay Lesbian Soc Serv.* 2009 October 1; 21(4): 357–373. doi:10.1080/10538720802498280.

## The Healthy Young Men's Study: Sampling Methods to Recruit a Random Cohort of Young Men Who Have Sex with Men

**Wesley L. Ford, MA, MPH [Director],**

Office of Ambulatory Care, Los Angeles County Department of Health Services, 313 N. Figueroa St, Suite 704, Los Angeles, CA 90012

**George Weiss [Senior Research Manager],**

Saban Research Institute, Childrens Hospital Los Angeles, 4650 Sunset Blvd, MS#30, Los Angeles, CA 90027

**Michele D. Kipke, PhD [Professor of Pediatrics and Preventive Medicine],**

Childrens Hospital Los Angeles and Keck School of Medicine, University of Southern California, 4650 Sunset Blvd, MS#30, Los Angeles, CA 90027

**Anamara Ritt-Olson, PhD [Senior Research Manager],**

Saban Research Institute, Childrens Hospital Los Angeles, 4650 Sunset Blvd, MS#30, Los Angeles, CA 90027

**Ellen Iverson, MPH [Assistant Professor of Pediatrics and Preventive Medicine], and**

Childrens Hospital Los Angeles and Keck School of Medicine, University of Southern California, 4650 Sunset Blvd, MS#30, Los Angeles, CA 90027

**Donna Lopez, MSW**

Saban Research Institute, Childrens Hospital Los Angeles, 4650 Sunset Blvd, MS#30, Los Angeles, CA 90027

Wesley L. Ford: wford@ladhs.org; George Weiss: gweiss@chla.usc.edu; Michele D. Kipke: mkipke@chla.usc.edu; Anamara Ritt-Olson: ritt@usc.edu; Ellen Iverson: eiverson@chla.usc.edu; Donna Lopez: dlopez@chla.usc.edu

### Abstract

Recruiting a scientifically sound cohort of young men who have sex with men (YMSM) is an enduring research challenge. The few cohort studies that have been conducted to date on YMSM have relied on non-probability sampling methods to construct their cohorts. While these studies have provided valuable information about HIV risk behaviors among YMSM, their generalizability to broader YMSM populations is limited.

In this paper the authors describe a venue-based sampling methodology used to recruit a large and diverse cohort of YMSM from public venues in Los Angeles County. Venue-based sampling is a multi-stage, probability sampling design that uses standard outreach techniques and standard survey methods to systematically enumerate, sample, and survey hard-to-reach populations. The study design allowed the authors to estimate individual, familial and interpersonal psychosocial factors associated with HIV risk and health seeking behaviors for a cohort of YMSM with known properties. Study participants completed an extensive baseline survey and over a two year period will complete four follow-up surveys at six-month intervals. The baseline survey was administered in both English and Spanish.

### Key words

Adolescents; men who have sex with men; gay; bisexual; hard-to-reach populations

## INTRODUCTION

Twenty-five years after the first diagnosed AIDS case, relatively little is known about ecological factors that influence risk or protective behaviors among young men who have sex with men (YMSM), and in particular, young men of color who have sex with men (Parker, Easton, and Klein 2000; Ross et al. 2004; Stokols 1992; Wohlfeiler 2000). Most surveys of this hard-to-reach population of YMSM have relied heavily on non-probability sampling methods such as respondent driven and convenience samples. While such studies have provided valuable information about ecological factors associated with risk and protective behaviors within this population, their findings have limited generalizability to the broader YMSM population.

Sampling hard-to-reach or hidden populations such as young men who have sex with men poses numerous practical and methodological challenges because the collection of accurate information about them is simply not possible using traditional sampling and estimation techniques. While not all YMSM are hidden, many of them may not want to be openly identified within their broader communities as a man who has sex with other men. Moreover, known sampling frames or lists of YMSM simply do not exist and, even if they did, a broad random sample from the overall population would not be cost-effective given their relatively small number in the overall population.

The few cohort studies that have been conducted on YMSM under 25 have relied on relatively small convenience or respondent driven samples focused exclusively on homeless populations (Meyer and Dean 1995; Rosario, Schrimshaw, and Hunter 2004; Whitbeck et al. 2004). To date, much of our current understanding of YMSM risk behavior come from two cross-sectional studies, the Young Men's Survey (YMS), conducted more than 10 years ago, and the Community Intervention Trial for Youth (CITY) project (Valleroy et al. 2000; Guenther-Grey et al. 2005). Both studies employed venue-based sampling techniques to select cross-sectional random samples of YMSM in metropolitan areas throughout the United States. The venue-based sampling technique used by these two studies has been adapted by Healthy Young Men's (HYM) study to recruit a cohort of YMSM.

HYM study is a longitudinal study that draws upon developmental theories of risk and resilience to examine the individual, familial, interpersonal, and community contexts in which drug use, HIV risk-related behaviors, and health promoting and seeking behaviors take place among YMSM. Venue-based sampling techniques similar to those used by YMS and CITY were used to recruit a racially and ethnically diverse cohort of YMSM. Cohort participants completed an extensive baseline interview and over the next two years will complete four follow-up interviews administered at six-month intervals. In this paper, we describe the venue-based sampling methods used to recruit the HYM study cohort. Descriptive statistics of the study cohort collected at baseline are also presented.

## STUDY ELIGIBILITY

Young men were eligible to participate in the study if they: 1) were 18 to 22 years of age, 2) self-identified as gay, bisexual or uncertain about their sexual orientation and/or they reported having had sex with men, 3) self-identified as Caucasian, African American, or Latino of Mexican descent, 4) were residents of Los Angeles County, and 5) anticipated living in Los Angeles County for the six months following recruitment. Young men of mixed race or ethnicity who identified with one or more of the targeted racial/ethnic groups were eligible to participate in the study. If they identified with two or more of the eligible racial/ethnic groups, they were asked to select the one they most identified with and were assigned to that specific

group. If they did not identify with one group more than another, they were randomly assigned to a group.

During the final three months of recruitment, the upper age limit was increased to 24 years of age for African American and 23 years of age for Caucasians. The upper age limit was increased for these two groups as the number of African Americans and Caucasians encountered in the 18–22 year old age group was smaller than that of their Latinos of Mexican descent counterparts. Raising the age limit ensured a sufficient number of young men would be recruited into the African American and Caucasian cohorts. A total of 32 young men (8 African American, N=126 and; 24 Caucasian, N=195) were enrolled into the study sample from these upper age groups.

These young men did not appear to differ from their younger counterparts on other demographic and key outcome variables in any significant way and, given their small number, we do not believe they introduced any significant known biases to the overall sample.

## STUDY METHODS

The venue-based sampling methodology was developed by YMS to estimate the prevalence of human immunodeficiency virus type I infection and related risk behaviors among YMSM, and was adapted by CITY to evaluate the effects of a multi-component, community level interventions aimed at promoting safe sex behaviors among YMSM 15–25 years of age (MacKellar et al. 1996; Muhib et al. 2001). Both YMS and CITY have demonstrated that venue-based sampling can be effectively used to sample a large and accessible population of YMSM to estimate HIV risk behaviors.

Venue-based sampling is a probability based sampling design that combines outreach techniques with standard survey methods to enumerate, sample, and survey hard-to-reach populations in public venues where they are known to congregate and socialize. Venues are enumerated during various times and days to construct a list of venues and their associated sampling periods. Only time periods determined to yield reasonable numbers of eligible respondents are included in a sampling frame and the remaining time periods are discarded. Venue-specific time periods are randomly selected each month from a sampling frame to construct monthly sampling calendars. In accordance with the sampling event calendars, young men who appear to be eligible are consecutively enumerated, approached, screened, and surveyed if determined eligible and they agreed to participate.

## HYM SAMPLING FRAME

During a six month formative research process, all public venues in Los Angeles County known to be frequented by young men who have sex with men were identified by reviewing local publications and Internet sites that cater to the gay community; conducting field observations and focus groups; briefly interviewing young men at various public venues; and interviewing key informant, service providers, and members of the target population. Venues such as bars, dance clubs, retail businesses, cafes and restaurants, health clubs, social and religious organizations, adult bookstores, bathhouses, high-traffic street locations, parks, beaches, and special events such as gay pride festivals, raves, and circuit and house parties were considered eligible for inclusion in the sampling frame. Whereas HIV positive support groups and clinical settings that routinely provide medical or mental health care or HIV/STD diagnostic or other prevention services were considered ineligible to prevent HIV positive YMSM from being overrepresented in the study cohort.

Staff contacted or met with venue owners and managers of commercial venues that were identified to explain the purpose of the study and to obtain permission to conduct study

activities on or near their premises. Organizations conducting outreach prevention or research activities at identified venues were also contacted to solicit their cooperation to avoid competing activities taking place during a HYM sampling event. All venue owners and prevention and/or research entities contacted cooperated fully with HYM staff during the recruitment phase of the study.

Young men were enumerated at different times of day at the venues identified during the formative research process to determine venue-specific sampling periods. Sampling periods are standard four-hour time periods on specific days that are expected to yield a minimum of eight eligible men. A venue may have more than one four-hour sampling period (for example, a bar had three sampling periods: Thursday 8:00 P.M. to 12:00 A.M., Friday 10:00 P.M. to 2:00 A.M., and Sunday 2:00 P.M. to 6:00 P.M.). Only venues with at least one sampling period were included in the sampling frame. The criterion of eight eligible men per sampling period reflects the number of young men that needed to be encountered at sampling events during the twelve-month recruitment period to enroll an adequate number of men into the cohort. To estimate whether eight eligible young men were expected to attend a venue during any given four-hour sampling period, two types of enumeration were conducted.

## TYPE I AND II ENUMERATIONS

Type I enumerations were conducted by a single study team member in high-traffic venues known to be frequented by large numbers of YMSM. During Type I enumerations, a study team member systematically counted young men who appeared to meet the study's age and race/ethnic eligibility criteria as they entered a predefined intercept area or crossed a predefined imaginary line during a 60-minute venue observation period. Young men who entered the area or crossed the line more than one time during an enumeration period were counted only once. A standardized enumeration formula, the product of the number of men counted by the length of time observed, standardized to a four-hour sampling period, was used to determine the potential number of eligible young men expected to attend the venue during a standard four-hour sampling period. Pilot enumerations found that approximately 50 percent of all young men who appeared eligible did not meet the study eligibility criteria when screened. Thus, if a Type I enumeration estimated that less than sixteen eligible young men were likely to attend a venue in a standard four-hour sampling period, a Type II enumeration was conducted.

Type II enumerations were also conducted when the formative research indicated the number of young men attending a venue was questionable or too small to warrant a Type I enumeration. Typically, Type II enumerations were conducted by two study team members working in tandem to count and interview a sample of young men. One team member - the enumerator - counted (clicked) all young men who appeared to meet the study eligibility criteria and entered a predefined intercept area or crossed an imaginary line at the venue during two 60-minute enumeration periods separated by a one hour time period. The second team member, using a standardized screening instrument, briefly screened a sample of the young men clicked by the enumerator to determine whether they met the study eligibility criteria. The screening instrument collected the basic demographic information needed to determine eligibility, i.e., age, date of birth, race/ethnicity, zip code, sexual identity, and the gender of all lifetime sexual partners.

No identifying information other than date of birth, which was used to verify the respondent's self-reported age, was collected on the screening instrument. It was not possible to link any of the information to a specific individual at a given venue. Participants were informed that they could refuse to respond to any question that made them uncomfortable.

Young men who entered the intercept area or crossed the intercept line more than one time during the enumeration period were counted only once. Attendance estimates were calculated

as the product of the number of young men counted and the number of young men determined eligible averaged for the two enumerations, and standardized to a four-hour sampling period. If a Type II enumeration estimated that eight or less eligible men were likely to attend the venue during a four-hour sampling period, two additional enumerations were conducted. When additional enumerations were conducted at a venue for the same sampling period, the number of young men counted and the number found to be eligible were averaged for all of the enumerations conducted at the venue to calculate the attendance estimate.

The attendance estimates derived from the Type I and II enumerations were used to construct a list of four-hour venue-day-time (VDT) sampling periods. For Type I enumerations, only VDTs that had attendance estimates of at least sixteen eligible men were included in the monthly sampling frame to take into account that approximately 50 percent of the young men counted would not meet the study eligibility criteria when screened. For Type II enumerations, only VDTs that had a minimum of eight eligible men were included. Forty-one (41) Type I and 47 Type II enumerations were conducted over a three-month period, using the enumeration procedures described above. The enumerations yielded a total of 80 VDTs that met the sampling frame inclusion criteria, and represented 36 different venues (Table 1).

### Special Events

The formative research phase identified special events that either occurred infrequently (e.g., gay pride events) and attracted large numbers of YMSM or occurred frequently on known days and times but the location was unpredictable (e.g., ball events and house parties). Ball events are lavish costume celebrations and competitions that act as social events for MSM of color. Because these special events were unique and YMSM who attended them may have differed from the young men who attended the other types of identified venues, two special VDTs were created to accommodate them in the sampling frame.

A “wildcard” VDT was created for events such as gay pride celebrations and health fairs that occurred only once during the 12 month recruitment period, and Ball Event and House Party VDTs were created for the ball and house party events that occurred several times a month on a known days and times but whose locations were unknown until the day of the event. Wildcards were placed on the calendar non-randomly. The Ball Event and House Party VDTs were included in the monthly sampling frame and were randomly selected with the other VDTs.

Because venues popular with youth may change rapidly the procedure for constructing the sampling frame was dynamic. Throughout the recruitment period participants were asked to identify public venues that they and their friends attended, and the study team routinely reviewed gay publications and Internet sites to identify new venues as they appeared in the community. As new venues were identified or existing venues became unproductive, they were enumerated or re-enumerated in accordance with the procedures described earlier. Based on the attendance estimates, VDTs were modified, added, dropped or re-introduced into the sampling frame each month before the monthly sampling calendars were constructed.

### MONTHLY SAMPLING CALENDARS

Each month, 16–24 VDTs were randomly selected with an equal probability from the updated sampling frame to create the monthly sampling calendars. In constructing the sampling calendar, wildcards for the month were first placed on the calendar and the remaining VDTs were placed on the calendar in the order of their random selection. For example, a VDT (Micky’s on Thursday 10 PM to 2 AM) was selected and placed on the calendar for the first available Thursday of the month. A second VDT was then selected and placed on the calendar in the same manner. If a VDT was randomly selected but it was not possible to schedule it on the calendar because all the available days for the selected VDT had already been scheduled,



it was discarded and another VDT was drawn. This process was repeated until all the monthly sampling events were scheduled.

Because unproductive sampling events were likely to occur for unforeseen reasons such as venue closures, unpredictable social mixing patterns, and inclement weather, alternative sampling events were selected for each primary VDT. Alternate VDTs were randomly selected from a list of all VDTs that occurred on the same day and time as the scheduled primary sampling event. In some instances no alternative sampling event existed for the primary sampling event. When a primary sampling event was found to be unproductive and no alternative sampling event existed, the primary sampling event was canceled and no sampling occurred for that event. A total of 215 sampling events were scheduled during the twelve month recruitment period, of which twelve were canceled due to some unforeseen reason such as inclement weather or because an event was unproductive and no alternative sampling event existed in the sampling frame.

## SAMPLING EVENTS

In accordance with the monthly sampling calendar, HYM staff counted and recruited cohort participants for a twelve-month period between February 2005 and January 2006. Each sampling event was conducted by a three to four person recruitment team with one team member designated the sampling event enumerator. Following the procedures described for Type II enumerations, young men who appeared to meet the study eligibility criteria were systematically counted (clicked) by an enumerator as they entered the venue's predefined intercept area or crossed a predefined imaginary line. After a young man was counted, the enumerator directed a recruiter to approach him to determine his eligibility using the eligibility screening instrument. Young men who stopped when approached were asked if they were willing to take a few minutes to answer some questions to assist with a research study. If a young man declined, he was thanked and the recruiter marked on a log the reason for refusal, if known. Young men who agreed to be screened were taken to a nearby location that offered some privacy to complete the eligibility screener that required approximately five minutes to administer. Individuals that did not meet the study's eligibility criteria, based on their responses to the screener questions, were given a small incentive and thanked for their time. Young men who met the eligibility criteria were provided an in-depth description of the study, including what commitments would be required of them if they agreed to participate. They were also given an opportunity to ask any questions they might have about the study. They were then asked if they were willing to be enrolled in the HYM cohort as a study participant, and if they agreed, the recruiter obtained their written, informed consent and their contact information. Contact information was collected on a separate form from that of the Informed Consent and screening instrument. The contact information form captured the participant's name, home, cell, work, and a friend's phone number(s), participant's e-mail address(es), and his home address and an alternative mailing address. A friend's phone number was collected for tracking purposes only and used as a back up, if the respondent could not be reached using the personal contact information he had provided.

Participants were also given an appointment date, time and location to complete the baseline questionnaire and were provided with information on how to contact the study team. If a participant was unable to commit to an appointment date and time at recruitment, a member of the study team contacted him within 24 hours to schedule a baseline interview date and time. All baseline interviews were scheduled within two weeks of the participant's recruitment date.

Consecutive recruitment of enumerated men who were determined eligible continued until all recruiters were occupied interviewing potential study participants. When all recruiters were busy interviewing, recruitment was temporarily suspended until such time that a recruiter

became available to approach the next consecutively enumerated young man. The enumerator continued to count men who appeared to meet the study's eligibility criteria while recruitment was temporarily suspended.

Young men were enumerated for the entire four-hour sampling period. The designated enumerator for the event served as the enumerator for the entire sampling period to minimize the chance of individuals being counted more than once. Eligible young men who declined to be enrolled in the study during a previous sampling event, and were randomly selected again at a subsequent sampling event, were eligible for recruitment into the study cohort. The screening and recruitment process was conducted in Spanish and English.

Of the 8,355 young men successfully approached during the 203 sampling events conducted during the twelve-month recruitment period, 56 percent (4,648) agreed to be screened, of which 1,371 (29%) met the study eligibility criteria. Of those eligible, 938 (68%) agreed to participate. A participant was not considered enrolled in the cohort until he completed the baseline interview. Of the 938 young men that agreed to participate in the study at the time of recruitment 526 (56%) completed a baseline interview and were enrolled in the study cohort.

The age, race/ethnicity, sexual identity, and sexual behavior of the eligible men who refused to participate, who agreed to participate but failed to keep their scheduled baseline interview, and those ultimately enrolled in the cohort were compared to determine if they were differentially distributed across the three groups on these four characteristics. While no differences were found for age, race/ethnicity, and sexual identity, the sampled young men who reported having had sex with men only were more likely to enroll in the cohort than those who reported having had sex with both men and women (Table 2). This difference may reflect the lack of specificity used to group the respondents by the biological sex of their lifetime sexual partners when screened at recruitment rather than differential selection into the cohort. For instance, young men whose lifetime sexual experiences were primarily with male partners, except for a single or limited number of experiences with female partners, were classified as having had sex with both men and women, as were those whose lifetime sexual experiences were primarily with female partners, except for a single or limited number experiences with male partners. These two groups may be qualitatively different from each other and, in turn, qualitatively different from those young men whose lifetime sexual partners were more or less equally distributed between males and females. Of the 220 cohort participants who reported both male and female lifetime sexual partners at recruitment, 83% reported that they had had sex with male partners only in the three month period prior to completing the baseline questionnaire, whereas none reported having had sex with female partners only, and 17% reported having had sex with both male and female partners. This suggests that the biological sex of these young men's sexual partners is highly skewed towards male partners and grouping them into a single group is likely to have introduced unknown biases.

## SURVEY INSTRUMENT

The baseline survey instrument assessed a wide range of constructs, including ethnic/national identity; educational background; work history; residential history; family history/composition; history of violence, including domestic violence; sources of social support; sexual orientation/identity; internalized homophobia; "connectedness" to community (ethnic/racial, geographic, work, gay); health and health care utilization, history of sexually transmitted infections, including HIV; HIV testing experiences; mental health concerns; alcohol and other drug use; and involvement in HIV risk and protective behaviors. The baseline survey consisted of validated measures, measures developed for use in previous research (Kipke et al. 1997), as well as new measures specifically designed from the study's formative research findings. The survey was then pilot tested with a sample of 51 YMSM who met the study criteria,

psychometric analyses were performed, and revisions were made to yield the final instrument used to collect the baseline data.

Participants were asked to report their age; race/ethnicity; place of birth, immigration status; current residence; employment status; whether they are attending school; whether they had ever been homeless; whether they had ever “exchanged a sexual act or favor for something like money, drugs, or a place to stay;” and whether they had ever participated in the street economy (e.g., selling/running drugs, prostitution, panhandling, theft).

Sexual orientation was assessed using three sources of information: 1) sexual identity, 2) sexual attraction to both males and females, and 3) disclosure of sexual identity/orientation. Sexual identity was measured by providing participants with a list of descriptors and asking which they most identified with (e.g., gay, queer, bisexual, same gender loving, down low, straight, heterosexual). Sexual attraction was measured by asking participants “How much are you sexually attracted to males/females” using two separate questions and a 5-point scale (1= very strongly, 2= somewhat strongly, 3= not very strongly, 4= not at all attracted to males/females, and 5= don’t know). The sexual attraction items were then recoded to create a new 3-level nominal variable: sexually attracted to males only, to females only, or to both males and females. Disclosure of sexual identity/orientation was measured by asking participants to report how many of their family, best/closest friends, other friends, classmates, coworkers, and teachers/school personnel know about their sexual identity or orientation, using a 4-point scale (1= all, 2= most, 3= some, 4= none).

## INTERVIEW PROCEDURES

The baseline questionnaire required approximately 1 1/2 hours to complete and study participants were reimbursed \$35 for their time after they completed the questionnaire. They were also given prevention materials and referral information for HIV prevention and other health or social services, if needed. When possible, the same individual that recruited the young man into the study administered the questionnaire, and the questionnaire was completed in the study office or a public location such as a coffee shop that was convenient for the participant. Prior to administration of the questionnaire, interviewers reviewed the informed consent, the purpose of the study, the participation requirements, and collected additional contact information beyond what was collected at the time of recruitment. The interviewer explained to the participant the importance of keeping his contact information current during the two-year study period, including the incentives he would receive for updating his personal contact information. Participants were asked to provide only personal information that they felt comfortable sharing.

## BASELINE SAMPLE CHARACTERISTICS

As summarized in Table 3, a total of 526 YMSM were enrolled in the study during the first twelve-months of recruitment, which included 195 (37%) Caucasian, 126 (24%) African American, and 205 (39%) Latino YMSM of Mexican decent. The average age was 20.1 years, with 40% of the sample being 18–19 years of age. Eighty-two respondents (16%) reported having been born outside of the US, while over half (54%) of the respondents reported living at home with their family at the time of their baseline interview. Twenty-two percent reported being in school while an additional 27% reported both attending school and being employed; only 13% reported being neither in school nor employed. Forty-three percent reported being very or somewhat religious, while 57% reported not being very religious or not at all religious. Of those that reported being religious, 71% reported that they were affiliated with the same religion that they had participated in while growing up, with the vast majority being Catholic (34%) or Protestant (33%).



Given the study enrollment criteria, 81% self-identified as homosexual, gay or some other same-sex sexual identity (e.g., “same gender loving”), 16% self-identified as bisexual, and 1% self-identified as straight or heterosexual. In contrast, 71% of the sample reported being sexually attracted exclusively to males, 27% reported being sexually attracted to both males and females, and 1% reported only being sexually attracted to females. Twenty-five percent reported having been previously diagnosed with an STI, with 6% having been diagnosed with two or more STIs. Three percent of respondents reported having tested positive for HIV, while 17% reported not knowing their HIV-status. Remarkably, 20% reported having a history of sexual abuse/assault, 16% reported having traded a sexual act or favor for something, and 9% reported a history of homelessness.

## DISCUSSION

The venue-based sampling methodology used by HYM allowed us to systematically recruit a large and diverse cohort of YMSM from venues where they are known to congregate and socialize in Los Angeles County. This approach presents a number of challenges. For instance, considerable effort is required to make certain that the maximum number of venues is included in the sampling frame. Sampling too few venues or only venues of one particular type may introduce unknown biases, as participants from a small set of venues or of one venue type may not be representative of all venues or venues of a single type. Additionally, the study’s methodological procedures for enumerating venues, constructing the sampling frame and monthly sampling calendars, and screening and recruiting study participants are labor intensive and must be strictly followed if a systematic sample is to be obtained.

As with any sampling methodology, venue-based sampling has its strengths and weaknesses. Because we sampled from a universe of VDTs, and not the overall population, not all members of the target population had an equal probability of being selected into the cohort. YMSM that did not attend any of the sampled VDTs had a zero probability of being enrolled in the cohort and those that rarely attended the sampled VDTs had a lower probability of being enrolled than those that regularly attended them. Thus, it is likely that some segments of YMSM were either under sample, or in the case of young men who do not identify as being gay, and therefore don’t attend gay venues, were not sampled at all.

A door-to-door survey in San Francisco found that 91% of gay and bisexual respondents aged 18–23 reported they had recently attended a gay bar (Catania 1998). These data support our assumption that by sampling public venues such as gay bars and clubs it is possible to sample significant segments, if not all segments, of the YMSM population.

Among our HYM participants, 89% reported they had attended a gay bar or club at least once in the past three months. Moreover, when we mapped our HYM participants by home zip code we found that they resided throughout Los Angeles County. These data suggest that we were able to recruit into our cohort a geographic and diverse cross-section of YMSM with known properties. It is unlikely that a comparable cohort of YMSM could have been recruited using any other method.

It is also important to note that it may not be feasible to use venue-based sampling in smaller cities in which the gay community is less visible and established, as the number and type of venues where YMSM are able to congregate and socialize in public may be limited. In small cities many YMSM may meet other young men through more informal networks.

Another potential weakness of our study is that it may not be possible to generalize our findings to YMSM who live outside of Los Angeles County. This would be particularly true if the underlying properties of the young gay men sampled in Los Angeles County differ significantly

from YMSM who live in other major urban areas in California or major urban areas in other parts of the country.

Moreover, a number of young men may rely solely on the Internet to socialize or to seek sex with other men; particularly if they do not identify themselves as being gay. Such young men represent an important and unique subset of YMSM. However, the venue-based sampling methodology we have described in this paper is not robust enough to capture these individuals and represents a limitation to venue-based sampling.

Venue-based sampling does, however, have a number of strengths compared to convenience samples. By collecting information on a participant's venue attendance patterns, enrollment probabilities can be estimated and weighted to adjust for YMSM who attend the venues more often and have a greater chance of being enrolled in the study. Using standard survey sampling methods, the weighted probabilities can then be used to estimate the prevalence of risk behaviors or psychosocial factors for young men who have sex with men that attend the venues included in the sampling frame. A detailed discussion of the formula used to weight data to correct for bias introduced by differential attendance patterns at venues is beyond the scope of this paper and readers are referred to MacKellar et al. (1996).

While our sampling methodology does not allow for inference to a geographically or demographically defined sample, it produced a broad random sample of YMSM whose properties can be characterized when making inferences to the larger YMSM population. Unlike respondent-driven sampling, this method does not depend on gaining entry to specific social networks. By systematically identifying and randomizing all feasible VDTs, consistently scheduling the sampling calendar, and systematically sampling individuals at sampling events, we are able to generalize our findings to a well defined and significantly broader segment of the YMSM population than would have been possible had we recruited our cohort through convenience and respondent-driven sampling methods.

It is clear that limited success has been achieved in preventing HIV infections among many young men who have sex with men as they continue to engage in high risk behaviors, despite decades of education and prevention efforts. It is equally clear that we need to better understand the psychosocial contexts and developmental processes in which risk and protective behaviors become instilled in each new generation of emerging adults. Given the central link between knowledge generation and knowledge utilization, it is important that researchers, practitioners, and policy makers understand the fundamental elements of research design if they are to construct and deliver effective social services to those they seek to serve.

Additionally, the emergence of this type of research, which is able to demonstrate the interaction and additive effect of various psychosocial characteristics associated with HIV risk and protective behaviors, will result in more refined prevention efforts. As such, research studies must explore new methods for constructing catholic, representative samples of young men who have sex with men. Venue-based sampling is one such method that can be used to recruit a broad cohort of YMSM, allowing us to expand our knowledge base and understanding of HIV among this hard-to-reach and vulnerable population.

## Acknowledgments

This study was funded by the National Institute on Drug Abuse of the National Institutes of Health (R01 DA015638-03). The authors wish to acknowledge the contributions of senior members of the research team, including Robert Paul Appleby, George Ayala, and Jennifer Unger, as well as members of the field research team, including Cesar Arauz-Cuadra, Marianne Burns, Julie Carpineto, Peter Cruz, Katrina Kubicek, Larkin Mackey, Bryce McDavitt, Miles McNeeley, Conor Schaye, Maral Shahanian, and Alex Tapia. The authors would also like to acknowledge the insightful and practical commentary of the members of The Community Advisory Board: Noel Alunit, Asian Pacific AIDS Intervention Team, Chi-Wai Au, LA County Dept of Health Services, Ivan Daniels III, Coordinator's Los Angeles

Black Pride, Ray Fernandez, AIDS Project Los Angeles, Trent Jackson, Youth/Trent Jackson Media Group, Dustin Kerrone, LA Gay and Lesbian Center, Miguel Martinez, Division of Adolescent Medicine, CHLA, Ariel Prodigy, West Coast Ballroom Scene, Brion Ramses, West Coast Ballroom Scene, Ricki Rosales, City of LA, AIDS Office, Haquami Sharpe, Minority AIDS Project

## References

- Catania, J. Urban Men's Health Study (UMHS) Methodology (unpublished report). San Francisco: University of California; 1998.
- Guenther-Grey C, Varnell S, Mathy R, Weiser J, O'Donnell L, Stueve A, Remafedi R. the Community Intervention Trial for Youth Study Team. Trends in sexual risk taking among urban young men who have sex with men 1999–2002. *Journal of National Medical Association* 2005;97(7):38–43. Supplement to July 2005.
- Kipke M, Montgomery S, Simon T, Iverson E. "Substance abuse" disorders among runaway and homeless youth. *Substance Use and Misuse* 1997;32(7 & 8):969–986. [PubMed: 9220564]
- MacKellar DA, Valleroy LA, Karon J, Lemp G, Janssen R. The young men's survey: Methods for estimating HIV seroprevalence and risk factors among young men who have sex with men. *Public Health Reports* 1996;111(Supplement 1):138–144. [PubMed: 8862170]
- Meyer H, Dean L. Patterns of sexual behavior and risk taking among young New York City gay men. *AIDS Education and Prevention* 1995;7(Supplement):13–23. [PubMed: 8664094]
- Muhib FB, Lin LS, Stueve A, Miller RL, Ford WL, Johnson WD, et al. A venue-based method for sampling hard-to-reach populations. *Public Health Reports* 2001;116(Supplement 1):216–222. [PubMed: 11889287]
- Parker R, Easton D, Klein C. Structural barriers and facilitators in HIV prevention: A review of international research. *AIDS* 2000;14(Supplement 1):S22–S32. [PubMed: 10981471]
- Rosario M, Schrimshaw EW, Hunter J. Ethnic/Racial Differences in the Coming-Out Process of Lesbian, Gay, and Bisexual Youths: A Comparison of Sexual Identity Development Over Time. *Cultural Diversity and Ethnic Minority Psychology* 2004;10(3):215–228. [PubMed: 15311975]
- Ross MW, Henry D, Freeman A, Caughy M, Dawson AG. Environmental influences on safer sex in young gay men: A situational presentation approach to measuring influences on sexual health. *Archives of Sexual Behavior* 2004;33(3):249–257. [PubMed: 15129043]
- Stokols D. Establishing and maintaining healthy environments: Toward a social ecology of health promotion. *American Psychologist* 1992;47:6–22. [PubMed: 1539925]
- Valleroy LA, MacKellar DA, Karon J, Rosen DH, McFarland W, Shehan DA, et al. HIV prevalence and associated risks in young men who have sex with men. *Journal of the American Medical Association* 2000;284(2):198–204. [PubMed: 10889593]
- Whitbeck L, Chen X, Hoyt D, Tyler K, Johnson K. Mental Disorder, Subsistence Strategies, and Victimization Among Gay, Lesbian, and Bisexual Homeless and Runaway Adolescents. *Journal of Sex Research* 2004;41(4):329–342. [PubMed: 15765273]
- Wohlfeiler D. Structural and environmental HIV prevention for gay and bisexual men. *AIDS* 2000;14(Supplement 1):S52–S56. [PubMed: 10981475]

**Table 1**  
Comparison of Study Eligibility Criteria for Total YMSM – Eligible to YMS Enrolled in the Study Cohort

Study Eligibility Criteria	Total Eligible	Refused study participation n (%)	Did not complete interview n (%)	Enrolled in cohort n (%)	Pearson Chi-Square
Age					
18–19	526	167 (31.7)	160 (30.4)	199 (37.8)	0.663
20–21	524	157 (30)	165 (31.5)	202 (38.5)	
22+	321	109 (34)	87 (27.1)	125 (38.9)	
Race/Ethnicity					
African American	355	108 (30.4)	121 (34.1)	126 (35.5)	0.24
Mexican Latino	528	178 (33.7)	145 (27.5)	205 (38.8)	
Caucasian	488	147 (30.1)	146 (29.9)	195 (40.0)	
Sexual Identity					
Gay/Bisexual/MSM	1353	426 (31.5)	406 (30)	521 (38.5)	0.92
Heterosexual	15	5 (33.3)	5 (33.3)	5 (33.3)	
Sexual Behavior					
MSM	557	158 (28.4)	156 (28)	243 (43.6)	0.01
MSMW	648	209 (32.3)	219 (33.8)	220 (34)	
Other	67	18 (29.9)	20 (29.9)	29 (43.3)	

**Table 2**

Description of the Study Sample (N=526)

Variables	Categories	n (%)
Age	18 – 19 yrs	206 (40)
	20 – 21 yrs	196 (37)
	22+ yrs	124 (24)
Race/ethnicity	African- American	126 (24)
	Caucasian	195 (37)
	Mexican descent	205 (39)
Immigration	Born in other country	82 (16)
Residence	Family	281 (54)
	Own place/apartment	191 (36)
	With friends/partner	36 (7)
	No regular place/other	17 (3)
Employment	In school	113 (22)
	In school, employed	142 (27)
	Employed, not in school	201 (38)
	Not employed, not in school	70 (13)
Sexual identity	Gay	391 (74)
	Other same-sex identity	38 (7)
	Bisexual	85 (16)
	Straight	3 (1)
	DK/RF	9 (2)
Sexual attraction	Males only	371 (71)
	Males and females	144 (27)
	Females only	6 (1)
	Neither, don't know, missing	5 (1)
Religiosity	Very religious	47 (9)
	Somewhat religious	180 (34)
	Not very religious	88 (17)
	Not religious	190 (36)
	Don't know	1 (0)
HIV Status	Positive	15 (3)
	Negative	420 (80)
	Don't know	90 (17)
STI (ever)		132 (25)
Sex exchange (ever)		85 (16)
Street economy (ever)		110 (21)
Homeless (ever)		45 (9)



**Table 3**

Venue type	Count
Social Service Agency	4
Ball/House Party	5
Bar/Club	18
Pride Festival	5
Street Corner	4