

Recruitment-Adjusted Estimates of HIV Prevalence and Risk among Men Who Have Sex with Men

Effects of Weighting Venue-Based Sampling Data

Samuel Jenness, MPH
HIV Epidemiology Program

APHA Annual Meeting
November 2010



Coauthors

New York City

Department of Health and Mental Hygiene

Christopher Murrill, PhD

Alan Neaigus, PhD

New York University

Center for Drug Use and HIV Research

Camila Gelpi-Acosta, MA

Holly Hagan, PhD

Travis Wendel, JD

Disclosures

The following personal financial relationships with commercial interests relevant to this presentation existed during the past 12 months:

No relationships to disclose

Venue-Based Sampling

- Roots in targeted sampling of hard-to-reach groups (Watters and Biernacki, 1989)
 - Populations ‘hidden’ from probabilistic sampling
- Used for target populations that congregate at definable and identifiable locations (‘venues’)
 - History in socio-behavioral HIV and substance use research on MSM and drug users
- All variations of VBS include elements of randomness in recruitment to improve upon convenience sampling in terms of external validity

VBS Example: CDC Young Men's Survey

- First large-scale use of VBS to conduct behavioral research
- Target population: sexually active MSM under 30 who attend MSM-oriented social venues in 7 U.S. cities
- Core VBS methods
 - Universal list of venues created
 - Venues randomly selected from list as recruitment event
 - Presumed MSM entering each venue sequentially and non-preferentially approached

Potential Recruitment Bias in VBS

- Validity of estimates for venue-attending target population influenced by two components of selection probability
 - Frequency of venue attendance: persons who attend venues more frequently have higher selection probability
 - Venue volume: persons who attend venues with low venue volume have higher *event-specific* selection probability
- Bias exists if outcome variables are associated with either
- Few VBS studies have accounted for this recruitment bias
 - Limited to one-component weighting and no direct comparison

Study Objectives

- Develop a dual-component weighting mechanism to adjust for the two potential recruitment biases
- Apply this mechanism to VBS data to examine the impact on outcomes
 - Core HIV risk factors related to venue characteristics
 - Prevalence of undiagnosed HIV infection

National HIV Behavioral Surveillance (NHBS)

- 21 cities throughout the United States
- Funded by CDC, designed collaboratively
- Ongoing, cyclical study of three risk groups: *MSM, IDU, and high-risk heterosexuals*
- NHBS-MSM data collection in 2008
- Cross-sectional study design
- Standardized, structured survey and HIV test

First-Stage Venue Sampling

- Universe of MSM-oriented venues constructed
 - Publications, ethnography, and observational research
 - Venues eligible where $\geq 75\%$ of venue patrons are adult MSM
 - Included commercial and non-commercial venues
 - Peak hours of operation (in four-hour blocks) also recorded
- All information placed in CDC-created randomizing software
 - Each month, list of randomly selected venues and corresponding time periods generated
 - Selections were placed on recruitment calendar with 1-2 non-random 'one-off venues'

Second-Stage Participant Sampling

- Field Recruitment Events
 - Mobile van at each venue during each event
 - All adult men who 'entered' the venue counted
 - Counted men sequentially approached to participate
 - Interested men were then screened for eligibility and all eligible men provide informed consent
- Eligibility criteria
 - Male gender, adult, New York City resident, English or Spanish speaking
 - Self-reported HIV+ and non-MSM eligible, but removed from this analysis

Response Weight Component

- Approach rate

MSM approached (MSM_a) / MSM counted at venue (MSM_c)

- Response rate

MSM interviewed (MSM_i) / MSM approached (MSM_a)

- Event-specific recruitment probability

MSM_i / MSM_c

- Response weight (W_r) reciprocal of recruitment probability

MSM_c / MSM_i

- W_r for HIV prevalence estimate calculated separately based on testing response rates

Attendance Weight Component

- Participants asked: “In the past 12 months, how often have you gone to a place where gay men hangout, meet or socialize?”
- Attendance weight (W_a) based on relative timeframe of response options

Response	Relative Frequency	Reciprocal (W_a)
1x/day	1 (reference)	1 (reference)
> 1x/week	3.5	1/3.5
1x/week	7	1/7
> 1x/month	15	1/15
1x/month	30	1/30
< 1x/month	60	1/60

Composite Weight and Analysis

- Composite weight (W_c) is the product of response and attendance weights: $W_r * W_a$
- For prevalence estimation comparisons, unweighted point estimates and 95% confidence intervals were compared with estimated weighted with W_c
- Unique recruitment events treated as dependent clusters controlled for when calculating standard errors
- Non-overlapping confidence intervals considered as conservative measure of statistical significance

Participant Zip Code of Residence and Location of Recruitment Venues

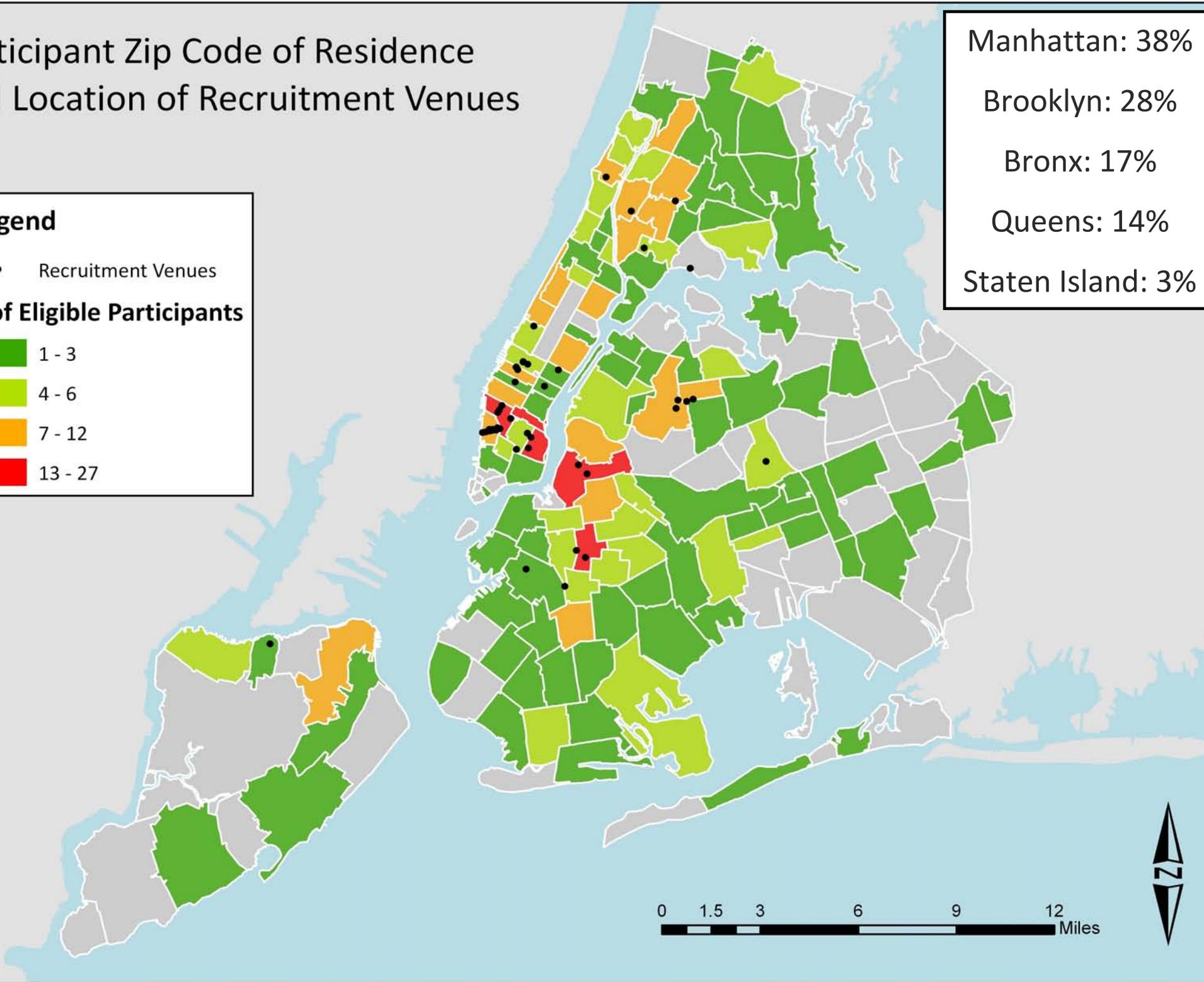
Manhattan: 38%
Brooklyn: 28%
Bronx: 17%
Queens: 14%
Staten Island: 3%

Legend

- Recruitment Venues

of Eligible Participants

Dark Green	1 - 3
Light Green	4 - 6
Orange	7 - 12
Red	13 - 27



Venue-Specific Approach and Response Rates

	Total		Venue Volume	Approach Rate	Response Rate
	n	%	mean	mean %	mean %
Overall	479	100	137.8	28.8	61.9
Recruitment Venue					
Bars	282	58.9	86.6	31.1	59.4
Clubs	37	7.7	173.4	16.1	53.4
Parks	39	8.1	39.6	44.8	69.8
CBOs	30	6.3	29.9	47.0	94.6
House Balls	29	6.1	308.7	15.5	71.8
Sex Strolls	31	6.5	227.9	11.9	67.6
Pride Events	31	6.5	538.9	14.7	38.6

Attendance-Specific Approach and Response Rates

	Total		Venue Volume	Approach Rate	Response Rate
	n	%	mean	mean %	mean %
Overall	479	100	137.8	28.8	61.9
Attendance Frequency					
1x/day	91	19.0	143.7	26.9	62.1
≥ 1x/week	144	30.1	100.7	31.0	60.4
1x/week	85	17.8	145.3	28.5	61.4
≥ 1x/month	60	12.5	144.7	28.0	59.3
1x/month	24	5.0	168.6	29.5	64.1
< 1x/month	75	15.7	177.9	27.6	66.6

Demographics

	n	Unweighted		Composite Weighting	
		%	95% CI	%	95% CI
Race/Ethnicity					
Black	125	26.1	22.1–30.0	43.0	27.8–58.2
Hispanic	167	34.9	30.6–39.1	33.5	24.3–42.8
White	152	31.7	27.5–35.9	19.7	8.0–31.3
Other	35	7.3	5.0–9.6	3.8	0.1–7.5
Age					
18-29	219	45.7	41.2–50.2	55.6	39.9–71.3
30-39	134	28.0	23.9–32.0	28.8	19.0–38.6
40-49	76	15.9	12.6–19.2	9.1	2.4–15.8
50+	50	10.4	7.7–13.2	6.5	1.3–11.6

Sexual Risk and Substance Use (Past Year)

	n	Unweighted		Composite Weighting	
		%	95% CI	%	95% CI
Unprot. Anal Intercourse	239	49.9	46.0–54.4	49.2	41.3–57.1
UAI with Cas/Exch Part.	102	21.3	17.6–25.0	9.0	2.8–15.3
≥ 5 Total Partners	210	43.8	39.4–48.3	25.9	15.5–36.4
Group Sex Encounters	85	17.7	14.3–21.2	9.3	5.1–13.6
Internet Sex Partners	136	28.4	24.3–32.4	29.6	23.3–35.9
Weekly Binge Drinking	120	25.1	21.2–28.9	11.0	4.7–17.2
Hard Drug Use	159	33.2	29.0–37.4	19.7	11.7–27.7

Undiagnosed HIV Infection

	n	Unweighted		Composite Weighting	
		%	95% CI	%	95% CI
HIV Serostatus					
Negative	326	82.5	78.8–86.3	85.5	78.0–92.9
Positive	69	17.5	13.7–19.9	14.8	7.4–22.2

Summary

- Developed a two-component weighting approach to adjust for differences in event-specific recruitment probability and overall venue attendance frequency
- Observed wide variation in these two components
- Adjusting for differential in these components resulted in:
 - Significantly higher prevalence estimates of young and black MSM
 - Significantly lower prevalence estimates of high-risk sexual and substance use behaviors
 - Marginally lower prevalence estimates of HIV infection

Recruitment Component

- Variation in recruitment probability partially depends on static size of study field team
 - Other research designs (PPS) vary the second-stage recruitment probability based on first-stage enumerations, but limited in social venue settings
- Biased estimates if outcome measures are related to venue characteristics
 - MSM sampled at parks and sex strolls are down-weighted as the approach and response rates were higher than average
 - May be responsible for lower weighted estimates of high-risk behaviors like UAI with casual/exchange partners

Attendance Component

- Differential attendance frequency at MSM-oriented social venues is natural to VBS design
 - NHBS survey question is proxy for overall patterns of attendance at all potential VBS venues
- Biased estimates if outcome measures are related to attendance frequency
 - Commercial venues directly or indirectly providing access to alcohol and drugs comprise majority of venues
 - Lower weighted estimates of binge alcohol and hard drug use may relate to down-weighting of high frequency bar patrons

Limitations

- These estimates and weighted estimates are not necessarily more valid than unweighted estimates
- No weighting component for first-stage selection probabilities; unknown multiplicities of venue attendance
 - Individual response weight unique to the event, and multiple opportunities for enumeration possible
- Weighting components may be influenced by seasonal variation and recall error
- Representativeness dependent on comprehensiveness and accuracy of venue universe

Conclusions

- Often no gold standard in estimating behavioral risks and disease outcomes in high-risk populations
 - VBS data particularly useful in designing prevention activities
- Like Respondent-Driven Sampling, ongoing development and evaluation of potential biases; work to address them in study design and analytically
- Future methodological research should focus on adjustments for participant selection probabilities across spectrum of first-stage sampling
- For now: standard reporting of response rates, recruitment by venue type, overall venue attendance frequency

Contact

Samuel Jenness, MPH

HIV Epidemiology Program

NYC Department of Health

Phone: 212-442-6445

Email: sjenness@health.nyc.gov