Analyzing Frequencies of Gay Men's Sexual Events Using A Generalized Poisson Regression Approach

K. P. Singh[1], J.T. Wulu[1, 2], A.A. Bartolucci[1], and T. Valappil[1]
[1]University of Alabama at Birmingham, Department of Biostatistics
Birmingham, AL 35294-0022, USA
[2]Camcare Health Education and Research Institute, CAMCARE, Inc.,
Charleston, West Virginia 25304, USA

Abstract: The determinants of gay men's frequencies of different sexual partners are examined among cohorts of gay men by applying generalized Poisson regression (GPR), Poisson regression (PR), and negative binomial regression (NBR) models. The models are estimated by the method of maximum likelihood. Approximate tests for the dispersion and goodness-of-fit measures for comparing the models are discussed. The data were collected from cohorts of gay men recruited in four AIDS Community Demonstration Project Cities (Dallas, TX; Denver, CO; Long Beach, CA; and Seattle, WA) from 1987 to 1991. Cohort participants shared information on HIV-risk, HIV-testing history, sexual activities, attitudes, smoking status, alcohol and drug use, trust for HIV prevention programs, and other social factors through questionnaires during first two visits. A total of 975 gay men showed complete data in the baseline and first follow-up study sample based on exclusion criteria. Of this sample, approximately 88% and 12% were white (non-Hispanic) and non-whites, respectively (age range: 18 to 62 years, median age: 31 years). About 24% reported high risk of being exposed to the AIDS virus; 33% indicated moderate risk; 38% adhered to using condoms consistently; and at least 50% had multiple male sexual partners and have had sex under the influence of alcohol and drugs. GPR analyses showed that all ages, never being married and categories of risk are inversely significantly related to the number of different sexual partners over time. Also, the effects of positive AIDS test results, condom use, and trust for AIDS prevention programs are negatively significantly related to the number of times gay men had sex under the influence of drugs and alcohol. HIV prevention and intervention efforts are needed and should focus on overcoming gay men's barriers to condom use and providing them the necessary information on risky sexual behaviors and safe sex practices.

1. INTRODUCTION

In June 1981, the Centers for Disease Control and Prevention (CDC) reported that between October 1980 and May 1981, five young gay men had been diagnosed with *Pneumocystis carinii pneumonia*, an illness that is currently known by its initials, PCP (CDC 1995b). In the next month, the CDC reported that 26 cases of the formerly very rare *Kaposi's sarcoma* (KS) had been diagnosed among gay men in the prior 30 months (CDC 1995b). These reports provided the first intimations of an epidemic that would infect very large numbers of gay men. Almost two decades later, the HIV/AIDS epidemic passed through the stage of being regarded as the "gay plague" or "gay-related immune disorder" to a disease affecting all populations regardless of individuals' sexual orientations, race, gender, nationality, or religious beliefs. Mann *et al.* (1992) reported that non-heterosexual, including gay men, transmission accounts for approximately 30% of all HIV infections worldwide.

Studies have indicated that the first half-million (actually 501,310) AIDS cases were reported in the United States by October 31, 1995 (CDC 1995a, b). This number represents close to one-half or more of all those estimated by the CDC to be infected with HIV by the end of 1989 (CDC 1990). Rosenberg (1995) estimated that the total number of persons ever infected with HIV was between 870,000 and 1,200,000 by the end of 1993. Levine *et al.* (1997) reported that based on such statistics, there were between 400,000 and 600,000 persons living with HIV, but not diagnosed with AIDS at that time. They further reported that of these half-million AIDS cases, 51% (256,000) were men who have sex with men (MSM) and 7% (35,000) had joint exposure from sex with men and injection drug use (IDU). Rosenberg (1995) claimed that persons who are infected with HIV are underreported in the AIDS registry by some 18%. Thus, the implication of this is that the number of cases in MSM category would be around 300,000 and in MSM/IDU group about 41,000. Levine *et al.* (1997) indicated that the exact number of gay men who are infected with HIV/AIDS might never be known. However, it is this number that most strongly expresses the role of the epidemic in transforming the communities of gay men of the HIV/AIDS generation and the successive generations of gay men.

Jonsen and Stryker (1993) revealed that by the late 1980s, the rate of new HIV/AIDS infections among gay men decreased, partly because those at highest risk had already been infected, but also because those still at risk had modified their behaviors. Other studies have shown that HIV/ AIDS cases among IDUs and heterosexuals have increased during the last decade and that MSM continue to account for the largest number of people reported with AIDS each year (CDC 1990, 1992, 1995c, 1999). In 1997 alone, 21,260 AIDS cases were reported among MSM, compared with 14,698 among IDUs and 8,112 among men and women who acquired HIV heterosexually (CDC-NCHSTP, 1998).

To account for the number of AIDS incidences, studies have examined the sexual behavior of MSM, the primary mode of transmission of AIDS (CDC 1999). Many gay or bisexual men would not be alive today had it not been for rigorous prevention efforts. Many members of the gay community find it overwhelming to be living in the midst of an epidemic and that cultural and spiritual issues often overcome the ability or desire to remain uninfected (Van Gorder 1995).

Dilley *et al.* (1998) reported in a cohort study of gay men that one in three of the participants in the study had lost a partner to AIDS and that a significant number of gay men, 25%, continued to engage in high risk sexual activity. Dilley *et al.* also reported that high-risk activity occurred among men who were knowledgeable about AIDS and understood the risk of their actions. They concluded that the men in the study sample engaged in these high-risk activities, regardless of whether their primary partner was himself HIV infected.

A CDC-sponsored study of gay men in San Francisco from 1994 to 1997 showed increases in unsafe sexual behavior among MSM in San Francisco, resulting in increased risk for HIV infection and transmission (CDC 1999). Among the study participants, it was found that one-third of MSM reported unprotected anal intercourse (UAI) with multiple partners during the previous six months. Additionally, substantial numbers of men interviewed in 1997 reported not knowing the HIV infection status of all their partners. Due to recent reports of increased rectal gonorrhea and the findings of this study, it is now suggested that there is an increased absence of safe sex behaviors found among men who have sex with men (CDC 1999).

In this study, participants are from four geographic locations in the United States: Seattle, Washington (32%); Denver, Colorado (21%); Dallas, Texas (29%); and Long Beach, California (18%). They were recruited as a result of mass advertisement in gay magazines and the use of posters boards in gay and lesbian communities. These efforts were facilitated via laboratories for investigation of behavioral changes, established by the National Centers for Prevention Services (NCPS) AIDS Community Demonstration projects and funded by the CDC. Two questionnaire instruments were developed and used to generate the data. The instruments are the initial (or baseline) visit questionnaire and the first follow up Visit Questionnaire. Trained staff conducted the interview with the study participants.

The data collection timelines ranged from November 30, 1987 to December 6, 1991. The original data included 4,244 initial respondents, with 2,087 at first follow-up. An overwhelming decreasing trend has been observed in subsequent follow-up visits because some participants died or got bored or disinterested due to tremendous exposure to prevention or educational HIV/AIDS programs in the late 1980s. In this study, 975 respondents were included at baseline and at first follow-up due to exclusion criteria. Records were excluded from the analysis if there was missing information for any of the following variables: c3501, c4001, and c5401. (See Table 1 for partial listings of variable definitions and Table 2 for descriptive statistics of selected variables.)

The study generated, for the first time, detailed nonlinear (e.g., PR, NBR, and GPR) analyses of such data. The models were used to assess the relationship between potential risk factors and HIV/AIDS status of a cohort of gay men. The main objective was to determine whether gay men modified their sexual behaviors between baseline and first follow-up time periods.

The dependent count variates included: number of different sexual partners and number of first-time sexual partners. The independent or explanatory variates included, among others, demographic variables (age, race/ethnicity, and marital status), behavioral risk, perceptions among gay men of HIV testing, exposure to risky sexual activities and drug/alcohol use, condom use, and seat belt use. Some explanatory variables were significantly correlated with others.

2. MODELS

Non-linear regression models (e.g., PR, NBR, and GPR) for assessing the relationships between the number of gay men's sexual partners and exposure to risky sexual behaviors are presented. These models are based on fundamental distributions: Poisson and generalized Poisson distributions.

2.1 POISSON REGRESSION MODEL

Suppose Y is a discrete random variable having independent response values $y_1, y_2, \ldots, y_n$ which follow a Poisson distribution . Then the Poisson regression (PR) model of Y given x_i may be defined as follows:

$$P(y_i; \underline{\beta}) = \frac{[\mu\{c_i(x_i,\beta)\}]^{y_i} e^{[-\mu(c_i(x_i,\beta))]}}{y_i!} \qquad (1)$$

where $i = 1, 2, \ldots, n$, $y_i = 0, 1, 2, \ldots$, and where $\underline{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$ is a p-dimensional vector of explanatory variables, $\underline{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ is a p-dimensional vector of unknown parameters, and c_i denotes some measure of exposure.

2.2 NEGATIVE BINOMIAL REGRESSION (NBR) MODEL

Let $Y_1, Y_2, \ldots, Y_n$ be a set of n independent random variables where Y_i follows a NBD with mean m_i and dispersion parameter k, denoted by $Y_i \sim NB(m_i, k)$. We have

$$P(Y_i = y_i) = \frac{\Gamma(k+y_i)}{y_i!\Gamma(k)}\left(\frac{k}{m_i+k}\right)^k\left(\frac{m_i}{m_i+k}\right)^{y_i} \qquad (2)$$

for $i = 1, 2, \ldots, n$ where $y_i = 0, 1, 2, \ldots$ and $m_i, k > 0$. The density function (2) is equivalent to

$$P(Y_i = y_i) = \frac{(k+y_i-1)!k^k m_i^{y_i}}{y_i!(k-1)!(k+m_i)^{k+y_i}} \qquad (3)$$

As $k \to \infty$, the NB converges in distribution to the Poisson. Let $\alpha = {}^1/_k$. When $\alpha = 0$, there is equi-dispersion; and when $\alpha > 0$, there is over-dispersion Consider the model:

$$\ln\ m_i = \beta_1 x_{i2} + \ldots + \beta_p x_{ip} \qquad (4)$$

where $x_{i_1}, \ldots, x_{i_p}$ are the values of p explanatory variables $x_1, \ldots, x_p$ which are associated with the random variable Y_i, and $\beta_1, \ldots, \beta p$ are p unknown parameters.

2.3 GENERALIZED POISSON REGRESSION (GPR) MODEL

The GPR model is a natural extension of the Poisson regression model.. Let Y be a generalized Poisson random variable depending on x_i. Observe that the expected value of the GPD is given by

$$\mu = \frac{6}{(1-\lambda)} = \theta\varphi, \qquad (5)$$

where $\varphi =$ $\frac{1}{(1-\lambda)}$ and $\theta = \frac{\mu}{\varphi}$, then the pdf of Y

$$P(Y=y|x_i) = \frac{\mu[\mu+(\varphi-1)y]^{y-1} e^{\frac{-[\mu+(\varphi-1)y]}{\varphi}}}{y!},$$

for $y=0,1,2,\ldots$

0, for $y > m$ when $\varphi < 1$ (6)

or otherwise where $\mu = \mu(\underline{x}) > 0$ and

$\varphi \geq \max(-{}^1/_2, 1-\frac{\mu}{4})$ denotes the square root of the index of dispersion and μ is the largest positive integer for which $\mu + m(\varphi - 1) > 0$, when $\mu < 1$. When $\varphi = 1$, the GPR model converges to the Poisson regression model; when $\varphi > 1$, the GPR model represents count data with over-dispersion; and when ${}^1/_2 < \varphi < 1$, the GPR model represents count data with under-dispersion provided $\mu > 2$. The GPR model is an outgrowth of both the GPD model and the Poisson regression model. Note that

$$\mu = \frac{6}{(1-\lambda)}, \qquad (7)$$

Consider the parametric transformation, $\lambda = \alpha\theta$. Thus,

$$\mu = \frac{6}{(1-\alpha\theta)} \qquad (8)$$

and then solve for θ. The new expression for θ shows that θ is a function of μ and α. Thus, we have a restricted GPR model with response variable, Y_i, a generalized Poisson random variable whose expected value, μ, is a function of the explanatory variables, x_i. The probability function of this restricted GPR model Y_i given x_i is defined by

$$f_i(y_i,\mu_i,\alpha) = \left[\frac{\mu_i}{1+\alpha\mu_i}\right]^{y_i}\left[\frac{(1+\alpha y_i)^{y_i-1}}{y_i!}\right]e^{\frac{-\mu_i(1+\alpha y_i)}{1+\alpha\mu_i}} \qquad (9)$$

where $y_i = 0, 1, 2, \ldots$, α is the dispersion parameter; when $\alpha > 0$, $V(Y_i|\underline{x}_i) > E(Y_i|\underline{x}_i)$ [i.e., over-dispersion]; when $\alpha < 0$, $V(Y_i|\underline{x}_i) < E(Y_i|\underline{x}_i)$ [i.e., under-dispersion]; when $\alpha = 0$, $V(Y_i|\underline{x}_i) = E(Y_i|\underline{x}_i)$ [i.e., equi-dispersion]; and $\mu_i = \mu_i(\underline{x}_i) = e^{\underline{x}_i'\beta_i}$, where $\underline{x}_i$ is a (k-1)-dimensional vector of explanatory variables and β_i is a k-dimensional vector of regression parameters.

The expected value of Y_i for any given x_i is defined by

$$E(Y_i|x_i) = \mu_i. \quad (10)$$

The variance of Y_i for any given x_i is defined by

$$V(Y_i|x_i) = \mu_i(1+\mu_i)^2. \quad (11)$$

3. RESULTS

With regards to demographic factors, over 80% of the respondents were White, single/never married, and related or acquainted with victims of HIV/AIDS. When analyzing behavior characteristics, it was found that over 97% and 99% of the gay men drank alcohol and used drugs, respectively. This is a clear indication of a high-risk group; however, approximately 75% rated themselves as moderate or low-risk for AIDS exposure. Also, nearly 75% of the respondents indicated they have fewer sex partners as a result of an interest in reducing risk for AIDS. Descriptive findings from the follow-up data indicate that 91% of the respondents were related or acquainted with HIV/AIDS victims; 95% reported using seat belts regularly. At follow-up, 97% of the respondents reported consistent (always) use of condoms during sexual activity, an increase from the baseline data that shows 50%. Additionally, 87% of the respondents reported they trust organizations of AIDS research and prevention programs.

Table 3 (baseline) shows significant and positive estimated dispersion parameters from NBR and GPR models. This is an indication that NBR or GPR models, rather than PR models, may be used to fit the data. From the goodness-of-fit measures in all models, GPR outperformed NBR and PR. In Table 3, respondents who rate themselves as high or moderate risk for AIDS exposure tend to have a significant positive effect on the number of sexual partners in the last three months. Further, respondents who report sexual encounters while under the influence of drugs or alcohol, binge drinking, or having tried drugs show significant positive contribution toward the number of sexual partners. Also, those who are regular smokers or heavy drinkers have a significant negative impact on the number of sexual partners. Respondents who have primary partner, engage in risky sexual activity with at least one one-time sex partner, interest in reducing risk of AIDS, consistent condom use during sexual activity, or believe/trust their partners are less likely to have more sexual partners. Significant positive effects on the number of sexual partners were found among respondents who have risky sexual activity with at least one occasional sex partner, or always used condoms during sexual activity, whether insertive or receptive. Respondents in the age groups 17-20, 20-30, or 30-40 were observed to have significant negative influences on the number of sex partners in year, prior to learning about AIDS.

Also, respondents who are single/never married have a significant negative impact on the number of sex partners. Respondents who rate themselves high-risk for AIDS exposure or are related/acquainted to HIV/AIDS victims had a significant positive effect on the number of sexual partners.

The relationship between the number of sexual partners and the need for sexual behavior modification in year prior to learning about AIDS was negative and significant. Respondents who have tried drugs were found to likely have multiple sexual partners. Also, respondents who indicate interest in reducing risk for AIDS were likely to have multiple sexual partners.

In the follow-up analyses, respondents who rate themselves high or moderate risk for AIDS exposure have a significant positive impact on the number of sexual partners in last three months. Also, respondents who have sexual encounters while under the influence of drugs/alcohol, have tried drugs, or believe their sexual partners had affairs contribute to an increase in the number of sexual partners in the last three months. Moreover, respondents who use seat belts regularly or those who trust in organizations of AIDS research and prevention programs tend to show significant negative effects on the number of sexual partners in last three months. Respondents who rate themselves as high or moderate risk respondents, have tried drugs, partner with IV drug users, or participate in sexual activity under the influence of drugs/alcohol show a positive relationship with the number of *first-time* sexual partners. Respondents who trust in organizations of AIDS research and prevention programs tend to show significant negative relationship with the number of *first-time* sexual partners.

The number of iterations (range: 3 to 19) for the PR, NBR, and GPR models from the SAS/ Fortran programs showed that the GPR models (range: 3 to 8) were competitive in generating non-linear solutions. More details of this data analysis are given in Wulu (1999).

4. CONCLUSION

The following covariates positively influence the frequency of male sexual partners for the cohort of gay men: drug-use, binge drinking, age, marital status, multiple sexual partners, and high/moderate risk individuals. Gay men who reported consistent condom use, smoking, heavy drinking, trust in HIV/AIDS prevention programs, fidelity with partner, and willingness to reduce AIDS cases are less likely to increase male sexual partners. In light of this, we recommend that HIV prevention and intervention efforts are needed and should focus on overcoming gay men's barriers to condom use and providing them the necessary information on risky sexual behaviors and safe sex practices.

5. ACKNOWLEDGEMENT

We acknowledge the staff of the CDC's NCPS AIDS Community Demonstration Projects for providing the data.

6. REFERENCES

CDC, HIV Seroprevalence estimates and AIDS Case projections in the United States: Report based on a workshop. *Morbidity and Mortality Weekly Report*, 39, 1-31, 1990.

CDC, NCP AIDS Community demonstration projects: What we have learned 1985-1990, prepared by Operational Research Section Be havioral and Prevention Research Branch, Di vision of STD/HIV Prevention, Center for Prevention Services. 1992.

CDC, *HIV/AIDS Surveillance Report* 7, 39. 1995a.

CDC, First 500,000 AIDS cases-United States, 1995. *Morbidity and Mortality* Weekly Report 44, 46. 1995b.

CDC, Deaths from melanoma-United States: 1973 to 1992. *Morbidity and Mortality Weekly Re port* 44, 337-347. 1995c.

CDC, Increases in unsafe sex and rectal gonorrhea among men who have sex with men-San Fran cisco, California, 1994-1997. *Morbidity and Mortality Weekly Report* 48, 45-48. 1999.

CDC-NCHSTP, Need for sustained HIV prevention among men who have sex with men: young and minority men at high risk. *HIV/AIDS Prevention Fact Sheet*. 1998.

Consul, P.C. Generalized Poisson Distribution: Properties and Applications. New York: Mar cel Dekker, Inc. 1989.

Consul, P.C. and Jain, G.C. A generalization of the Poisson Distribution. *Technometrics* 15, 791-799. 1973.

Dilley, J.W., McFarland, W., Sullivan, P. and Dicepola, M. Psychosocial correlates of un protected anal sex in a cohort of gay men attending and HIV-negative support group. *AIDS Education and Prevention* 10, 317-326, 1998.

Jonsen, A.R. and Stryker, J. eds. The social impact of AIDS in the United States. Washington. Washington D.C.: National Academy Press. 1993.

Levine, M.P., Nardi, P.M. and Gagnon, J.H. In changing Times: Gay men and lesbian encounter HIV/AIDS. Chicago, IL: The University of Chi cago Press.

Rosenberg, P.S. The scope of the AIDS epidemic in the United States. *Science* 270, 1372-1375. 1995.

Schnell, D.J., Galavotti, C. and O'Reilly, K.R. An evaluation of sexual behavior change using statistical and cognitive models. *Statistics in medicine* 12, 219-227, 1993.

Stokes, J.P., Vanable, P. and McKirnan, D.J. Com paring gay and bisexual men on sexual behavior, condom use, and psychosocial variables related to HIV/AIDS. *Archives and Sexual Behavior* 26, 383-396, 1997.

Van Gorder, D. Building community and culture are essential to succeeeful HIV prevention for gay and bisexual men. *AIDS and Public Policy Journal*. 1995.

Wulu, J.T. Generalized Poisson Regression Models with Applications, Ph.D. Thesis (Unpublished). University of Alabama-Birmingham, USA. 1999.

Table 1. Gay men's sexual behavior variable definition (short list) .

Variable	Definition
Baseline Dependent/Response Variables	
C3501	Number of male sexual partners in the last three months
C5401	Number of male sexual partners in year prior to learning about AIDS
Baseline Independent/Predictors or Covariates	
C4001	Number of sexual encounters while under the influence of drugs or alcohol
AGE1	Respondents between 17 and 20 years of age
AGE2	Respondents between 20 and 30 years of age
AGE3	Respondents between 30 and 40 years of age
RACE1	White respondents
MAR1	Single/never married respondents
RISK1	Respondents who rated themselves as high-risk for AIDS exposure
RISK2	Respondents who rated themselves as moderate risk for AIDS exposure
RISK3	Respondents who rated themselves as low-risk for AIDS exposure
REL_AIDS	Respondents who know or are related to HIV/AIDS victim(s)

Table 2. Gay men's sexual behavior descriptive statistics (short list)

(Sample Size=975)

Variable	Proportion of 1's	Mean	Std Dev(SD)	Min	Max
Baseline Dependent/Response					
C3501		4.1559	5.2686	0	75
C5401		15.0359	32.3393	0	350
Baseline Independent/Predictors or Covariates					
C4001		4.6051	8.4050	0	85
AGE1	0.033				
AGE2	0.518				
AGE3	0.352				
RACE1	0.878				
MAR1	0.858				
RISK1	0.236				
RISK2	0.334				
RISK3	0.401				
REL_AIDS	0.858				

Table 3. The effects of demographic covariates on the number of male sexual partners gay men had in last 3 months: Poisson, negative binomial, generalized Poisson regression using baseline data.

Variable	Poisson Regression			Negative Binomial Regression			Generalized Poisson Regression		
	Estimate	Standard Error (SE)	Wald t-statistics	Estimate	Standard Error (SE)	Wald t-statistics	Estimate	Standard Error (SE)	Wald t-statistics
Intercept	0.9910	0.1425	6.9561*	0.9924	0.2367	4.1933*	0.9918	0.2420	4.0982*
AGE1	-0.1469	0.0994	-1.4778	-0.0653	0.1835	-0.3556	-0.0376	0.1945	-0.1931
AGE2	-0.3039	0.0545	-5.5787*	-0.2610	0.1012	-2.5806*	-0.2473	0.1060	-2.3327
AGE3	-0.0663	0.0538	-1.2316	-0.0436	0.1019	-0.4282	-0.0369	0.1074	-0.3438
RACE1	0.0906	0.0501	1.8070*	0.0965	0.0893	1.0810	0.0980	0.0924	1.0602
MAR1	0.0055	0.0465	0.1193	-0.0383	0.0855	-0.4481	-0.0538	0.0904	-0.5952
RISK1	0.7352	0.1207	6.0931*	0.7174	0.1918	3.7411*	0.7157	0.1936	3.6960*
RISK2	0.6014	0.1203	5.0005*	0.5801	0.1899	3.0548*	0.5763	0.1911	3.0156*
RISK3	0.2375	0.1208	1.9651*	0.2246	0.1898	1.1837	0.2229	0.1905	1.1700
REL_AIDS	0.0376	0.0480	0.7849	0.0557	0.0843	0.6615	0.0609	0.0869	0.7005
Dispersion Parameter:									
α				0.5436	0.0333	16.3090*	0.2114	0.0111	19.0155*
Pearson's Chi-Square		5422.0668						5455.0489	
Generalized Chi-Square					1563.2930			1412.7067	
Deviance		3489.6400			939.5419			840.8362	
Log-Likelihood		-3163.2535			-2396.2320			-2364.3818	
Number of Iterations:		5			5			6	

*Significant at 0.05 level.

Note: See Table 1 for variable definitions.