The Lack of Selection Bias in a Snowball Sampled Case-Control Study on Drug Abuse

CLAUDIA S LOPES,* LAURA C RODRIGUES** AND ROSELY SICHIERI*

Lopes C S (Instituto de Medicina Social, State University of Rio de Janeiro, R. Sao Francisco Xavier 524, 7 andar, 20559–900 Rio de Janeiro–RJ, Brazil), Rodrigues L C and Sichieri R. The lack of selection bias in a snowball sampled case-control study on drug abuse. *International Journal of Epidemiology* 1996; **25**: 1267–1270.

Background. Friend controls in matched case-control studies can be a potential source of selection bias based on the assumption that friends are more likely to share exposure factors.This study evaluates the role of selection bias in a case-control study that used the snowball sampling method based on friendship for the selection of cases and controls.

Methods. The cases selected for the study were drug abusers located in the community. Exposure was defined by the presence of at least one psychiatric diagnosis. Psychiatric and drug abuse/dependence diagnoses were made according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R) criteria. Cases and controls were matched on sex, age and friendship. The measurement of selection bias was made through the comparison of the proportion of exposed controls selected by exposed cases (p1) with the proportion of exposed controls selected by unexposed cases (p2). If p1 = p2, then, selection bias should not occur.

Results. The observed distribution of the 185 matched pairs having at least one psychiatric disorder showed a p1 value of 0.52 and a p2 value of 0.51, indicating no selection bias in this study.

Conclusion. Our findings support the idea that the use of friend controls can produce a valid basis for a case-control study.

Keywords: Case-control studies, snowball sampling, selection bias, friendship matching, drug abuse

Epidemiological research into drug abuse has increased during the last three decades. The recognition of the drug problem as a worldwide phenomenon has emphasized the need to identify a drug user profile, risk factors for use, and the consequences of such use. However, the ascertainment of untreated drug abusers by means of traditional sampling methodologies has often shown poor results. The illegality of drug usage and the heterogeneity of drug users make representative community survey studies difficult. Hence most of the epidemiological studies carried out in this area have been conducted in clinical populations.[1-5] These studies contribute to the overall understanding of the phenomenon, but are likely to be limited in their generalizability. The difficulty of studying hidden populations reveals an opposition between extensive survey methodologies and intensive data collection. Surveys cannot reach hidden populations without specific adaptations and intensive data collection methodology is not obtained without the loss of easy generalizability.[6]

Bias is particularly acute in the study of hidden populations. The High School Senior Survey and the Household Survey are examples. Both surveys were sponsored by the National Institute on Drug Abuse, and had the primary purpose of measuring the prevalence and correlates of drug use in the USA.[7,8] These studies have been heavily criticized because their strategy for data collection (extensive survey) made it possible to miss entire hidden populations, resulting in inaccurate estimations of drug use. The High School Senior Survey contains data from only those seniors present on the day the survey was administered. The approximate 20% of each cohort that become school dropouts and those who are frequently absent are more likely to use drugs and are not adequately represented in these surveys.[9] Similarly, many drug abusers are homeless and are not represented by the Household Survey.[10] Indeed, the President's Commission on Organized Crime[11] recommended oversampling higher school dropouts and people without residence in these surveys. They also suggested that the snowball sampling method was

* Instituto de Medicina Social, State University of Rio de Janeiro, R. Sao Francisco Xavier 524, 7 andar, 20559–900 Rio de Janeiro–RJ, Brazil.
** London School of Hygiene and Tropical Medicine, University of London, UK.

appropriate for the generation of intensive results in this area.

The snowball technique is a method that yields a sample based on referrals made by people who share or know others who present the characteristics that are of research interest.[12] Historically, this method has been widely used in qualitative studies of hidden populations. The identification of such populations requires a knowledge of insiders who can locate people willing to participate in the study, and this method appears to be particularly applicable when the focus of interest is an area of deviant or illegal behaviour, as in the case of drug abuse.

In this study, the snowball technique seemed to be particularly useful because it allowed us to combat the problem of selection bias by matching because study participants were automatically matched on friendship. Matching has been used to minimize selection bias in studies in which the population base has not been precisely defined or because there is no accurate way of sampling it.[13] In these situations, matching on friendship or neighbourhood, which represent a number of undefinable sociodemographic factors that are impossible to quantify, increases comparability and thus, decreases the likelihood of selection bias.[14]

One of the concerns of this study was the possibility of selection bias as a consequence of friendship matching. This would occur if people who present a psychiatric disorder are likely to name a friend with the same psychiatric characteristics. In this paper we evaluated the possibility of selection bias in a friendship matched case-control study about psychiatric risk factors for drug abuse which used the snowball technique for the selection of the study sample.[15]

METHODS

The data were collected from 15 January to 15 July 1992, from 370 adults aged 18–40 who lived in Rio de Janeiro, Brazil. Six trained interviewers administered the Composite International Diagnostic Interview (CIDI)[16] to obtain information on demographic factors, history of psychiatric disorders and drug abuse/dependence. Computer diagnoses were generated according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R) criteria.[17] Exposure was defined as the presence of at least one psychiatric diagnosis and outcome by the presence of a drug abuse/dependence diagnosis.

The cases in the study were drug abusers. Both cases and controls were identified through the chain referral or snowball technique. Twelve drug abusers were initially identified with the help of ex-drug abusers, treatment seeking drug abusers and counsellors located at a drug abuse treatment facility and research centre, attached to the State University of Rio de Janeiro. All agreed with the interview, and each one was asked to name a friend who was a drug abuser (a new case) and a friend who had never been involved in the abuse of drugs (control). This chain referral was continued as new cases were asked to nominate friends until 201 cases were identified. None refused to participate. Of these 201 cases, 16 were excluded, 15 because they did not fulfil the criteria for a drug disorder and one because he was not able to name a friend control. All the elegible controls were included in the study. During recruitment, if a nominated control was later found to be a case, then this respondent was accepted as a case and two new controls were recruited. No control refused to participate.

In order to avoid loss of information due to the mental status of the respondent and poor recall of age of onset of disease, exclusion criteria included: 1) severe evidence of cognitive impairment in language or communication, 2) being aged under 18 or over 40, and 3) having a history of recent psychiatric or drug treatment.

The measurement of selection bias was made through the comparison of the proportion of controls with at least one psychiatric disorder selected by cases with at least one psychiatric disorder (p1) with the proportion of controls with at least one psychiatric disorder selected by cases without psychiatric disorders (p2). According to Flanders and Austin,[18] selection bias does not occur if p1 = p2.

RESULTS

Sixty five per cent of the cases and 52% of the controls were classified as having had at least one lifetime psychiatric disorder. The observed distribution of the matched 185 pairs with or without at least one psychiatric disorder is given in Table 1. The proportion of exposed controls selected by exposed cases (p1) was 0.52 and the proportion of exposed controls selected by unexposed cases (p2) was 0.51, indicating no selection bias in this study.

DISCUSSION

Selection bias can be a consequence of friendship matching if selection of subjects is not independent of the exposure under study within each stratum of the matching factors.[18] If the prevalence of exposure among the friends of exposed cases (p1) is equal to the prevalence of exposure among the friends of unexposed cases (p2), then bias resulting from the use of friend

TABLE 1 *The observed distribution of at least one psychiatric disorder among 185 case-control matched pairs, selected through the snowball technique. Rio de Janeiro, Brazil, 1992*

Cases	Controls		Total
	With at least one disorder (Exposed)	Without disorder (Unexposed)	
With at least one disorder (Exposed)	63 (A)	57 (B)	120 (A+B)
Without disorder (Unexposed)	33 (C)	32 (D)	65 (C+D)
Total	96 (A+C)	89 (B+D)	185

Where:

A = Number of matched pairs in which both the case and control are exposed.

B = Number of matched pairs in which the case is exposed and the control unexposed.

C = Number of matched pairs in which the case is unexposed and the control is exposed.

D = Number of matched pairs in which both the case and control are unexposed.

$$p1 = \frac{A}{A+B} \quad \text{and} \quad p2 = \frac{C}{C+D}$$

controls did not occur.[19] This follows from Flanders and Austin's formulation;[18] at least for the situation in which it is not necessary to stratify the matched pairs. A comparison of p1 and p2 may not be useful if pairs are matched on a correlate of exposure, as is a confounder. In this situation the matched pairs would be correlated, p1 and p2 would differ and selection bias introduced. In this study, the main objective of using the snowball technique was the sampling of a hidden population. In such a situation, matched pairs need not be correlated in terms of exposure. They would be correlated only if the matching variable, in this case friendship, was a true confounder.

The finding of no selection bias corroborates the conclusions drawn by Flanders and Austin and Pike and Robins[18,19] that selection bias should not arise if cases and their friend controls belong to the same subpopulation, i.e. to the same stratum of a population and therefore p1 = p2.

ACKNOWLEDGEMENT

REFERENCES

[1] Rounsaville B J, Weissman M M, Kleber H. Heterogeneity of psychiatric diagnosis in treated opiate addicts. *Arch Gen Psychiatry* 1982; **39:** 161–66.

[2] Ross H E, Glazer F B, Germanson T. The prevalence of psychiatric disorders in patients with alcohol and other drug problems. *Arch Gen Psychiatry* 1988; **45:** 1023–31.

[3] Kleinman P H, Miller A B, Millman R B *et al.* Psychopathology among cocaine abusers entering treatment. *J Nerv Ment Dis* 1990; **178:** 442–47.

[4] Bukstein O G, Glancy L J, Kaminer Y. Patterns of affective comorbidity in a clinical population of dually diagnosed adolescent substance abusers. *J Am Acad Child Adolesc Psychiatry* 1992; **31:** 1041–45.

[5] Carroll K M, Rounsaville B J, Bryant K J. Alcoholism in treatment-seeking cocaine abusers: Clinical and prognostic significance. *J Stud Alcohol* 1993; **54:** 199–208.

[6] Meter K M. *Methodological and design issues: techniques for assessing the representatives of snowball samples.* In: *The Collection and Interpretation of Data from Hidden Populations.* Rockville: NIDA Research Monograph, 1990.

[7] Johnston L, O'Malley P, Bachman J. *National Trends in Drug Use and Related Factors Among American High School Students and Young Adults, 1975–1986.* Washington, DC: DHHS Publication No. ADM87-1535, US Government Printing Office, 1987.

[8] National Institute on Drug Abuse. *National Household Survey on Drug Abuse: Main Findings 1985.* Washington, DC: DHHS Publication No. ADM88-1586, US Government Printing Office, 1987.

[9] Clayton R R, Ritter C. The epidemiology of alcohol and drug abuse among adolescents. *Adv Alcohol Subst Abuse* 1985; **4:** 69–97.

[10] Johnston L D, O'Malley P M. Why do the nation's students use drugs and alcohol? Self-reported reasons from nine national surveys. *J Drug Issues* 1986; **16:** 29–66.

[11] President's Commission on Organized Crime. *Report to the President and the Attorney General. America's Drug Habits: Drug Abuse, Drug Trafficking, and Organized Crime.* Washington, DC: Superintendent of Documentation, US Government Printing Office, 1986.

[12] Biernacki P, Waldford D. Snowball sampling: Problems and techniques of chain referral sampling. *Sociol Methods Research* 1981; **2:** 141–63.

[13] Clayton D, Hills M. *Statistical Models in Epidemiology.* Oxford: Oxford University Press, 1993.

[14] Kleinbaum D G, Kupper L L, Morgenstern H. *Epidemiologic Research. Principles and Quantitative Methods*. New York: Van Nostrand Reinhold, 1982.

[15] Lopes C S, Lewis G, Mann A. Psychiatric and alcohol disorders as risk factors for drug abuse. A case-control study among adults in Rio de Janeiro, Brazil. *Soc Psychiatry Psychiatr Epidemiol* 1996; **31:** (In press).

[16] Robins L N, Wing J, Wittchen H U *et al.* The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988; **45:** 1069–77.

[17] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association, 1987.

[18] Flanders W D, Austin H. Possibility of selection bias in matched case-control studies using friend controls. *Am J Epidemiol* 1986; **124:** 150–53.

[19] Pike M, Robins J. Re: 'Possibility of selection bias in matched case-control studies using friend controls' (Letter). *Am J Epidemiol* 1989; **130:** 209–10.

(Revised version received May 1996)