Catalogue no. 11-522-XIE

## Statistics Canada International Symposium Series - Proceedings

# Symposium 2004: Innovative Methods for Surveying Difficult-to-reach Populations

2004

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2004
Innovative Methods for Surveying Difficult-to-reach Populations

# ESTIMATING THE SIZE OF IDU POPULATION USING NEEDLE EXCHANGE PROGRAMS

Shenghai Zhang, Ping Yan and Chris Archibald [1]

## ABSTRACT

This article describes a sampling and estimation scheme for estimating the size of injecting drug use (IDU) population using Needle Exchange Programs (NEPs). It is designed to use the information of number of needles distributed in the NEPs centres. The approach involves a sampling design to collect a sample of injecting drug users (IDUs) who appear at NEPs in a certain period of time and to obtain retrospective self-report data on the number of friends among the IDUs and number of needles exchanged for each sampled injecting drug user. A methodology is developed that estimates the size of injecting drug users who have ever used the NEPs during the fixed period of time, and which allows us to estimate the proportion of injecting drug users in the group of using NEPs. The size of the IDU population is estimated by dividing the total number of IDU in group of using NEPs during the period of time by the estimated proportion of IDUs in the group. The technique hold promise for providing data needed to answer questions such as “How many members of the population in Canada are IDU?” and “Is that number changing?” and better understanding the dynamics of IDU population in a national level.

KEYWORDS: Estimation; IDU Population; Needle Exchange Programs.

## 1. INTRODUCTION

Injection drug use plays an important role in the global HIV/AIDS epidemic. In 2003, HIV infection associated with IDU transmission has been reported in more than 130 countries. In some countries, IDU is now the main mode of transmission of HIV. It is roughly estimated that worldwide in 2003 there are more than 13 million injecting drug users and in some regions more than 50% of them are infected with HIV (UNAIDS, 2004). The negative health consequences of IDU are not limited to just HIV infection. Sharing injection equipment carries a high risk of transmission of other blood-borne infectious diseases such as hepatitis B and hepatitis C. Also, injection drug use contributes to the epidemic's spread far beyond the circle of those who inject. Injection drug users, their partners, and their children account for at least 36% of all AIDS cases reported in the U.S. through 1999 (CDC, 2001). In Canada, injection drug use is also a problematic activity. Although it is difficult to obtain accurate data on the prevalence and profile of injection drug users in Canada, it is clear that there are large numbers of IDUs across the country (Achibald et al., 2001, Weekes and Cumberland, 2004).

Policy makers and researchers have realized that data on the size and pattern of drug injecting need to be systematically collected for a comprehensive understanding of the HIV epidemics among IDUs and of the efforts to prevent (Dehne, Adelekan, Chatterjee and Weiler, 2002). Understanding something about the dynamics of the injection drug users makes it possible not only to assess the likely impact of the spread of HIV/AIDS and other related diseases, but also alert policy makers to a worsening situation, or alternatively to provide evidence that prevention and other initiatives may be working. However, for most cites, countries and regions worldwide, no reliable estimates of the number of injecting drug users exist (Dehne, Adelekan, Chatterjee and Weiler, 2002), and those figures relating to the extent and pattern of drug injecting and their trends that are available are often the "guesstimates" of health officials rather than the results of research or systematic surveillance. However, policy makers and researchers have met the problem of collecting accurate information about the size and behaviour of injecting drug users, since they are not easily captured in a general population based survey (a typical generation population based survey is to survey individuals in a random sample of households in a district, province or a

[1]Centre for Infectious Disease Prevention and Control, Public Health Agency of Canada, Ottawa (email:szhang@hc-sc.gc.ca)

country, depending on the scale of the study). Injecting drug could be viewed as socially unacceptable. Therefore, injection drug users belong to the hard-to-reach group. Injectors sometimes even hide their habit from those with whom they live, including parents, roommates, and sexual partners. Recently, public health researchers, statisticians and sociologists have worked on the issue of this limitation of current methodology.

## 2. ESTIMATING THE SIZE OF IDU POPULATION WHO USE NEPS

Optimal approaches to disease prevention require regular and accurate estimates of disease prevalence. Needle and syringe exchange programs, which sterile needles and syringes are free or at a minimal cost for injecting drug users, are a convenient means of monitoring the prevalence of bloodborne viral infections among large numbers of injecting drug users who are currently injecting drugs. There are well over two hundred NEPs in Canada, with more under development (Health Canada, 2004). In addition, there are now numerous pharmacies that provide needle exchange services (Health Canada, 2004). For example, in 2002, a total of 1,109,056 needles were exchanged through the NEPs in Regina of Canada (Regina 2002 (addictions) report, 2004).

Our purpose in this Section is to estimate the size of IDU population who using NEPs, based on information of the number of needles distributed in a fixed period of time. To obtain such information from NEP centres, a simple selection of NEP centres to be interviewed may not be appropriate. To estimate the size of IDU population who use the NEP, we propose an alternative two stage-samples: we first select a sample of NEP centres, then, in a survey week, all individuals in the selected NEP centers are observed, which allows for the sampling of individuals at a limited number of centres.

How are the NEP centers chosen? Because of time and cost constraints, considering characteristics of IDU population, a sampling strategy has to be developed to select NEP centres. They could have been chosen completely at random from a list of the all centres of NEPs. However, to decrease the variability of parameter estimators based on data from the completed survey, sampling theory suggests that it is better to sample large centres (covering large-IDU-population) with higher probability than small centers (covering small IDU-population). So, the sampling design for NEP centers we consider is quite general and frequently used, that is the Stratified Probability-Proportional-to-Size sampling of NEP centers. Based on the number of needles which has exchanged in each NEP centre in the last year, we can divide all centers across Canada into $L$ strata with $K_h$ needle exchange centers in the $h$th stratum. In addition, needles are legally available for purchase through pharmacies (Miller and Tyndall, et al., 2002), although the willingness of pharmacists to sell syringes to IDUs is variable. This kind of exchange sites can be one of strata. From the $h$th stratum, $k_h$ NEP centers are sampled with inclusion probabilities $\pi_{h1}$, $\pi_{h2}, \cdots,$ $\pi_{h,k_h}$. The inclusion probabilities are determined by the number of needles that were exchanged in the last year. The $\pi_{hi}$ could be determined in the following way. Assuming that sampling is without replacement, the inclusion probability of the $hi$th NEP center in the list of $k_h$ NEPs centers is $\pi_{hi} = k_h Q_{hi} / \sum_{j=1}^{K_h} Q_{hj}$, where $Q_{hj}$ is the total needles that were exchanged at the $i$th centre in $h$th strata in the last year. It is necessary to require $\pi_{hi} < 1$. This is the case for all $i$, when $k_h = 1$. It could be expected that there is no very large $T_{hj}$ within each stratum. So, the requirement can be satisfied. Otherwise, we could set $\pi_{hi} = 1$ for the $i$ such that $k_h Q_{hi} / \sum_{j=1}^{K_h} Q_{hj} > 1$ (Sarndal, Swensson and Wretman, 1992).

It is clear that the size of the IDU population which a NEP center covers is correlated with the number of needles distributed. Therefore, Probability-Proportional-Size sampling offers the possibility of decreased variability of estimators of totals (Cochran, 1977). It is also offer some operational advantages in multistage sampling in that it can equalize the workload across geographic areas sampled at the first stage of sampling.

The needle exchange centres will be selected on the basis of large number of clients and representation from all Canadian jurisdictions. At present, there is a survey, I-track survey, is being conducted by Health Canada for describing the prevalence of HIV infection and HIV related risk behaviour among injecting drug users in Canada. The question, "how many needles did you use in the last three months?" could be added to the questionnaire. All injecting drug users attending the selected needle exchange centres during a specific week will be asked to complete a brief questionnaire and provide a venous blood sample that is tested for HIV and HCV.

Let $T_{hi}$ be the number of needles that are exchanged during the period of the three months in the $i$ sampled NEP centre of $h$ th stratum. Suppose there are $N_{hi}$ IDUs who were exchanged the needles in the last three months in this centres and $N_{hi_j}$ is the number of IDUs who has used $j$ needles in the last three month and all those needles come from this centre. The $N_{hi}$ and $N_{hi_j}$ include not only IDUs who attend the NEP centre but also those IDUs who use the needles from the centre through second exchange. A survey (DesJarlais, McKnight, Eigo and Friedmann, 2002) which conducted in American shows that 90% of NEPs actively encouraged secondary exchange which is defined as "providing needles that you know will be used by persons other than the exchanger". Therefore, the number of clients for a NEP centre dose not equal to the number of IDUs who exchange needles from this centre. The method provided above can estimate the size of IDU population who use NEP more accurately.

Now we can get relation between the number of IDUs who use needles from the centre and the total number of needles that are exchanged in the centre for the period of the three months,

$$T_{hi} = \sum_{j=1}^{m} jN_{hi_j} = N_{hi} \sum_{j=1}^{m} jR_{hi_j}$$

where the $m$ is the maximum number of needles that an IDU could be used in the three month, the $R_{hi_j} = N_{hi_j} / N_{hi}$ which can be estimated by the sample obtained in the survey week. Notice that $N_{hi}$ and $N_{hi_j}$ are unknown. However, what we need to know is the proportion $R_{hi_j}$ and it can be estimated by the sample proportion $r_{hi_j} = n_{hi_j} / n_{hi}$, where the $n_{hi}$ is the number of injecting drug users who attend the needle exchanges in the exchange centre and complete the questionnaire in the survey week and among of them there are $n_{hi_j}$ IDUs who used $j$ needles in the last three month. Therefore, the $N_{hi}$ can be estimated by

$$\hat{N}_{hi} = \frac{T_{hi}}{\sum_{j=1}^{m} jr_{hi_j}} .$$

In fact, the nominator $r_{hi} = \sum_{j=1}^{m} jr_{hi_j}$ is the mean of numbers of needles that exchanged for IDUs who used the NEP for their changes in the period of three months. The estimator of the size $N$ of IDU population who used NEPs at least once in the last three months in all Canadian jurisdictions is the weighted mean

$$\hat{N} = \sum_{h=1}^{L} \sum_{i=1}^{k_h} \pi_{hi}^{-1} \hat{N}_{hi} .$$

Suppose the IDUs who use the NEP centre have the same patterns as the survey week. In another wards, we assume that the sample of IDUs who are observed in the survey week is obtained by a simple random sampling without

replacement in the IDUs who use the NEP centre in the period of the three months. From the surveys which were carried out for studying the risk behaviour of IDUs in Australian (MacDonald, Wodak and et. Al., 1997) and Canada (Peggy, Myers, Calzavara and et. Al., 2003), the IDUs observed from the survey week are representative for the IDUs who use the NEP centres in the period of the three months. So, this assumption about the sample seems reasonable. Algorithms for Stratified Probability Proportion-to-Size Sampling selection without replacement could be realized, based on many methods, such as Hanurav-Vijayan algorithm (Vijayan, 1968) and Brewer's algorithm (Brewer, 1963).

## 3. ESTIMATING THE PROPORTION OF IDU POPULATION IN GROUP OF NEP USERS

Under the assumption in last Section, the IDU population is made up of two groups of people based on their status of participating NEPs. To estimate the size of IDU population, we have to know not only the size of the population who use NEPs but also the proportion of the population who participate the NEPs during the three months. The idea from respondent-driven sampling (Salganik and Heckathorn, 2004) will be used to estimate the proportion.

The sampled IDUs in the survey week in the $hi$th centre can be divided into two groups: one is for having used NEPs at least one time in the period of the three months and the other one is for having never used NEPs in the period, the groups are denoted by $\Omega_{NEP}^{hi}$ and $\Gamma_{\overline{NEP}}^{hi}$, respectively. The size of the group $\Gamma_{\overline{NEP}}^{hi}$ usually is small. To increase the size and coverage, we can use each IDU in the group of $\Gamma_{\overline{NEP}}^{hi}$ as an initial seed to conduct a respondent-driven sample, that is, each seed will recruit certain number of IDUs and provide information how many friends they have with NEPs users group and non-users group respectively. The new sample collected by members of the group $\Gamma_{\overline{NEP}}^{hi}$ will contain NEP users and non-NEP users. The $\Omega_{\overline{NEP}}^{hi}$ denotes the group which include all sampled non-NEP users. Now, to estimate the population proportion, we need to know the networking structure, such as the average degree of friendships, and probability that a non-NEP user (or NEP user) have a friendship with a NEP user.

The total number of friendships radiating from the IDUs in the group $\Omega_{NEP}^{hi}$ is denoted by $\Phi_{NEP}^{hi}$ which can be written as

$$\Phi_{NEP}^{hi} = n_{NEP}^{hi} \Psi_{NEP}^{hi},$$

where $\Psi_{NEP}^{hi}$ is the average friendships of IDUs in the group and $n_{NEP}^{hi}$ is the number of IDUs in the group. Let $\Pi_{NEP,\overline{NEP}}^{hi}$ be the number of friendships that all IDUs in group $\Omega_{NEP}^{hi}$ have with IDUs who don't use NEPs. Then the probability $P_{NEP,\overline{NEP}}$ that an IDU who uses NEPs has a friendship with an IDU who has never used the NEPs in the period of the three months can be estimated by

$$\hat{P}_{NEP,\overline{NEP}} = \frac{\sum_{h=1}^{L} \sum_{i=1}^{k_h} \Pi_{NEP,\overline{NEP}}^{hi}}{\sum_{h=1}^{L} \sum_{i=1}^{k_h} \Phi_{NEP}^{hi}}.$$

The average number of friendships of IDUs in the group $\Omega_{NEP}$ can be estimated by pooling the samples from each sampled centre together:

$$\hat{\Psi}_{NEP} = \sum_{d=1} d \frac{\sum_{h=1}^{L} \sum_{i=1}^{k_h} f_{NEP}^{hi}(d)}{\sum_{h=1}^{L} \sum_{i=1}^{k_h} n_{NEP}^{hi}},$$

where the $f_{NEP}^{hi}(d)$ is the number of IDUs in the group $\Omega_{NEP}^{hi}$ who have number $d$ friendships among IDUs. Similarly, we can get $\hat{P}_{\overline{NEP},NEP}$ and $\hat{\Psi}_{\overline{NEP}}$.

Based on the status of an IDU using NEPs in the period of the three months, we say that an IDU in state $S_{NEP}$ if he or she has ever used NEPs in this period, otherwise we say he or she is in state $S_{\overline{NEP}}$. Suppose we have respondent-driven sampling design to collect a sample. We will say the initial IDU (a seed) is chosen in step 0. Another IDU could be chosen based on the degree of friendships of the initial IDU. We say this IDU is chosen in step 1, and so on. Suppose the chance of recruiting another IDU with state $S_{NEP}$ depends on this chosen IDU only through his or her degree of friendships among IDUs. Suppose also that if the IDU chosen in step 0 is in state $S_{NEP}$, then an IDU in state $S_{\overline{NEP}}$ will be chosen with probability $1-P_{NEP,\overline{NEP}}$; and if the IDU chosen in step 0 is not in state $S_{NEP}$, then an IDU in state $S_{NEP}$ with probability $P_{\overline{NEP},NEP}$ chosen in step 1. Letting $X_n$ denote the status of an IDU chosen in the $n$th step, then $\{X_n, n=0,1,\cdots\}$ is a two-state Markov chain having a following transition probability matrix.

$$\begin{pmatrix} P_{NEP,NEP} & P_{NEP,\overline{NEP}} \\ P_{\overline{NEP},NEP} & P_{\overline{NEP},\overline{NEP}} \end{pmatrix}$$

It is clear that it is an irreducible ergotic Markov chain (Ross, 1997). In fact, we assume that IDUs within the group of using NEPs have similar proportions of degree of friendships with group of IDUs not using NEPs (or using NEPs). Also, notice that $P_{NEP,\overline{NEP}} = 1-P_{NEP,NEP}$ and $P_{\overline{NEP},\overline{NEP}} = 1-P_{\overline{NEP},NEP}$. In practice, it is much easier for a sample is collected by 1 wave. So, we have to use the different method to approach the proportions $\pi_{NEP}$ and $\pi_{\overline{NEP}}$ of a chain of IDUs in states $S_{NEP}$ and $S_{\overline{NEP}}$ respectively. Based on the Markov theory, we have

$$\pi_{NEP} = \frac{P_{\overline{NEP},NEP}}{1+P_{\overline{NEP},NEP}-P_{NEP,NEP}} \text{ and } \pi_{\overline{NEP}} = \frac{1-P_{NEP,NEP}}{1+P_{\overline{NEP},NEP}-P_{NEP,NEP}}.$$

The $\pi_{NEP}$ and $\pi_{\overline{NEP}}$ can be estimated by plug-in estimates $\hat{P}_{NEP,NEP}$ and $\hat{P}_{\overline{NEP},NEP}$. Now, based on the result given by Salganik and Heckathorn (Salganik and Heckathorn, 2004), the proportion $\alpha$ of IDUs who have ever used NEPs in the period of the three months can be estimated by

$$\hat{\alpha} = \frac{\hat{\pi}_{NEP}\hat{\Psi}_{\overline{NEP}}}{\hat{\pi}_{\overline{NEP}}\hat{\Psi}_{NEP}+\hat{\pi}_{NEP}\hat{\Psi}_{\overline{NEP}}}.$$

Now that we have estimated the size $\hat{N}$ of IDU population who use NEPs and the proportion $\hat{\alpha}$ of IDUs in using NEPs among all IDUs, we can use this information to estimate the size of IDU population as $\hat{N}/\hat{\alpha}$.

## 4. CONCLUSIONS

For many years researchers have tried to get accurate size of IDU population. We have shown that this goal could be fulfilled by the sample design combining harm reduction programs -- NEPs. Switching to this harm reduction programs and using the information of number of needles exchanged in the NEP centres give us a fresh and novel

approach to the estimation of the size of IDU population. Using NEPs allows us to design a sampling and estimation scheme which is both cheaper and more accurate than existing methods commonly in use and the methodology described here could be quicker and easier to implement compared to other methods for estimating the size of IDU population.

The NEPs provide a comprehensive HIV and blood-born infections prevention model to prevent the further spread of the diseases among IDUs and they have been proved to be effective for intervention of risk behaviour among IDUs. It is possible that the network of NEPs could embrace a well distributed, age and sex representative population of the area where we are interested in. However, it may be difficult to choose all NEPs to participate the data collection system. For the purpose of estimating the size of IDUs using NEPs, the sample design we proposed follows the basic principle of sampling theory which is that each individual in the target population should have some nonzero chance of being sampled in the survey (Korn and Graubard,1999). Notice that the target population in our first stage of estimation is all IDUs who have ever used the NEPs in the period of the three months. Therefore, each IDUs using NEPs has some nonzero chance of being sampled in the survey of estimating the size of IDUs of participating NEPs.

We use the idea of respondent-driven sampling to estimate the proportion of IDU population in using NEPs. However, in order to estimate the proportion, what we need to know is transition probability that a non-NEP user (or NEP user) has a friendship with a NEP user. Then the results from Markov chain and the article (Salganik and Heckathorn, 2004) were used for the estimates.

We have provided estimators which corresponding to the sample design. It is well known that ignoring the sample design can produce misleading estimates, especially, when a sample is selected via non-random sample design and then to ignore the sample design and treat the samples as if it was a simple random sample.

Our approach has concentrated on estimating the size of IDU population. It may be necessary to combine this sampling and estimation strategy with study of risk behaviour of IDUs. Further study for variances of the estimators has to be carried out. We have presented a number of analytic results and these analytic arguments could be further supported with numerical simulation which have not done yet. Also, the possible bias exists when we estimate the transition probabilities, because the sample size of non-NEP users is usually small. However, it may be possible to predict the magnitudes and direction of this bias ahead of time.

We hope that application of the techniques described here are possible for estimating the size of IDU population in Canada.

## APPENDIX

We discuss the variance of the estimator $\hat{N}$ in this appendix. First, we look at the estimator $\hat{N}_{hi}$. Let $Y_{hi}$ denote the number of needles exchanged by an injecting drug user in the centre for a period of the three months. The probability mass function of $Y_{hi} = j$ is denoted by $f_{hi}(y)$. The mean of $Y_{hi}$ is $\overline{Y}_{hi} = \sum_j jf_{hi}(j)$ . The variance of $Y_{hi}$ conditioned on $N_{hi}$ is $\operatorname{var}(Y_{hi}) = \frac{N_{hi}}{N_{hi}-1}\left[\sum_{j=1}^{N_{hi}} j^2 f_{hi}(j) - \left(\sum_{j=1}^{N_{hi}} jf_{hi}(j)\right)^2\right]$. The $N_{hi}$ and in $f_{hi}(j)$ could be substituted by $\hat{N}_{hi}$ and $r_{hi_j}$, respectively; the $\frac{N_{hi}}{N_{hi}-1}$ can be replaced by 1 because of $N_{hi}$ is usually very large in our case. Under these discussions and the notice that $T_{hi}$ is known and $\overline{Y}_{hi}$ is estimated by $r_{hi}$, we use linearization methods (Thompson, 1997 and Lohr, 1999) to approximate the variance of the estimator $\hat{N}_{hi}$ and

we have $E\left[\left(\hat{N}_{hi}-N_{hi}\right)^2\right]\approx\frac{1}{\hat{Y}_{hi}^2}N_{hi}^2\operatorname{var}\left(\bar{y}_{hi}\right)$, where $E$ denote the expectation. Finally, the variance $\hat{N}_{hi}$ can be estimated by

$\hat{\operatorname{var}}\left(\hat{N}_{hi}\right)=\frac{1}{\bar{y}_{hi}^2}\hat{N}_{hi}^2\left(1-\frac{n_{hi}}{\hat{N}_{hi}}\right)\frac{\left(\hat{\operatorname{var}}\left(Y_{hi}\right)\right)^2}{n_{hi}}$, where $\hat{\operatorname{var}}\left(Y_{hi}\right)=\frac{\hat{N}_{hi}}{\hat{N}_{hi}-1}\left[\sum_{j=1}^{\hat{N}_{hi}}j^2r_{hi_j}-\left(\sum_{j=1}^{\hat{N}_{hi}}jr_{hi_j}\right)^2\right]$. Because the sampling is carried in each stratum independently and within a stratum the Probability Proportional-to-size sampling with fixed size without-replacement design is used, the variance of the estimator $\hat{N}$ for the size of IDU who use the NEP in the period of the three months can be written as the following (Sarndal, Swensson and Wretman, 1992).

$$\operatorname{var}\left(\hat{N}\setminus\hat{N}_{hi},i=1,\ldots,K_h\right)=\frac{1}{2}\sum_{h=1}^{L}\sum_{i=1}^{K_h}\sum_{l=1}^{K_h}\left(\pi_{hi}\pi_{hl}-\pi_{hil}\right)\left(\frac{\hat{N}_{hi}}{\pi_{hi}}-\frac{\hat{N}_{hl}}{\pi_{hl}}\right)^2$$

where $\pi_{hil}$ is the covariance between indicator variables $I_{hi}$ and $I_{hl}$ ($I_{hi}=1$ if the centre $hi$ is included in the sample, 0 otherwise). For the fixed size without replacement design, the calculation of $\pi_{hil}$ could be complicated. For example, if the sample size in $h$th stratum is 2, the $\pi_{hil}$ has the following expression under the scheme proposed later (Sarndal, Swensson and Wretman,1992): $\pi_{hil}=\frac{2Q_{hi}Q_{hl}(Q_h-Q_{hi}-Q_{hl})}{Q_h(\sum_{j=1}^{K_h}c_{hj})(Q_h-2Q_{hi})(Q_h-2Q_{hl})}$,

where the $Q_h=\sum_{j=1}^{K_h}Q_{hj}$ and $c_{hj}=Q_{hj}(Q_h-Q_{hj})/(Q_h(Q_h-2Q_{hj}))$. The scheme is that, for the first draw, give the center $hi$ the probability $p_{hi}=c_{hi}/\sum_{j=1}^{K_h}c_{hj}$ of being selected; without replacing the first drawn element, (say $h,i_0$), and give the other element the probability $p_{hi|h,i_0}=Q_{hi}/(Q_h-Q_{h,i_0})$. The variance $\hat{\operatorname{var}}\left(\hat{N}\right)$ can be calculated by

$$\operatorname{var}\left(\hat{N}\right)=E\left(\operatorname{var}\left(\hat{N}\setminus\hat{N}_{hi},i=1,\ldots,k_h\right)\right)+\operatorname{var}\left(E\left(\hat{N}\setminus\hat{N}_{hi},i=1,\ldots,k_h\right)\right).$$

Considering that the $\hat{N}_{hi}$'s are estimated independently, we could have

$$\operatorname{var}\left(E\left(\hat{N}\setminus\hat{N}_{hi},i=1,\ldots,k_h\right)\right)\approx\sum_{h=1}^{L}\frac{K_h}{k_h}\sum_{i=1}^{k_h}\operatorname{var}\left(\hat{N}_{hi}\right).$$

By the Taylor linearization technique, we have an approximation of the variance of the estimator $\hat{N}$ for the size of injecting drug users who use NEP in the period of the three months is

$$\hat{\operatorname{var}}\left(\hat{N}\right)=\sum_{h=1}^{L}\left[\frac{1}{2}\sum_{i=1}^{k_h}\sum_{l=1}^{k_h}\left(\pi_{hi}\pi_{hl}/\pi_{hil}-1\right)\left(\frac{\hat{N}_{hi}}{\pi_{hi}}-\frac{\hat{N}_{hl}}{\pi_{hl}}\right)^2+\frac{K_h}{k_h}\sum_{i=1}^{k_h}\hat{\operatorname{var}}\left(\hat{N}_{hi}\right)\right].$$

## ACKNOWLEDGMENT

Authors thank Professor Mary E. Thompson at University of Waterloo and Professor J. N. K. Rao at Carleton University for useful suggestions and comments. The authors are also grateful to the members of HAMEG (HIV/AIDS Modeling and Estimating Group) of Centre for Infectious Disease Prevention and Control of Public Health Agency of Canada for the discussion.

## REFERENCES

Addictions Services (2004), "Regina 2002 Report", *A Canadian community epidemiology network on drug use report,* Regina.

Archibald, C. P., Jayaraman, G. C., Major, C., Patrick, D. M., Houston, S. M. and Sutherland, D. (2001), "Estimating the size of hand-to-reach population: a novel method using HIV testing data compared to other methods", AIDS, 15, pp. s41-s48.

Brewer, K. W. R. (1963), "A model of systematic sampling with unequal probabilities", *Australian Journal of Statistics*, 5, pp. 93-105.

CDC (2001), "HIV Prevention Strategic Plan Through 2005", Centers for Disease Control and Prevention, USA.

Cochran, W. G. (1997), "*Sampling Techniques*", 3rd Edition, New York: Wiley.

Dehne, K. L., Adelekan, M., Chatterjee, A. and Weiler, G. (2002), "The need for a global understanding of epidemiological data to inform human immunodeficiency virus (HIV) prevention among injecting drug users", LIV, *Bulletin on Narcotics,* pp.117-130.

DesJarlais, D. C., McKnight, C., Eigo, K. and Friedmann, P. (2000), "2000 United States Syringe Exchange Program Survey", The baron Edmond de Rothschild chemical dependency institute, New York.

Health Canada (2004), "Harm reduction and injection drug use: an international comparative study of contextual factors influencing the development and implementation of relevant policies and programs", http://www.hc-sc.gc.ca/

Korn, E. L. and Graubard, B. I. (1999), "*Analysis of Health Surveys*", Wiley series in probability and statistics survey methodology section, New York: John Wiley & Sons.

Lohr, S. L. (1999), "*Sampling: Design and Analysis*", New York: Duxbury Press.

MacDonald, M., Wodak, A. D., Ali, R. et al. (1997), "HIV prevalence and risk behaviour in needle exchange attendees: a national study", *Med. J. Aust*., 166, pp. 237-240.

Miller, C. L., Tyndall, M., Spittal, P., Li, K., Ralepu, A. and Schechter, M. T. (2002), "Risk-taking behaviours among injecting drug users who obtain syringes from pharmacies, fixed sites, and mobile van needle exchanges", *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 79, pp. 257-265.

Peggy, M., Myers, T., Calzavara, L., Wallace, E., Major, C. and Degani, N. (2003), "Regional variation in HIV prevalence and risk behaviours in Ontario injection drug users", *Canadian Journal of Public Health*, 94, pp. 431-435.

Roos, S. M. (1997), "*Probability Models*", 6th Ed. London: Academic Press.

Salganik, M. J. and Heckathorn, D. D. (2004), "Sampling and Estimation in Hidden Population using Respondent Driven Sampling", *Sociological Methodology*.

Sarndal, C., Swensson, B. and Wretman, J. (1992), "*Model Assisted Survey Sampling*", Series in Statistics, New York: Springer-Springer.

Thompson, M. E. (1997), "*Theory of sample surveys*", London: Chapman & Hall.

UNAIDS (2004), “2004 Report on the global AIDS epidemic – Executive Summary”, Joint United Nation Programme on HIV/AIDS, Geneva, Switzerland.

Vijayan, K.(1968), “An exact $\pi_{ps}$ sampling scheme: Generalization of a method of Hanurav”, *Journal of the Royal Statistical Society*, 54, pp. 499-513.

Weekes, J. and Cumberland, K. (2004), “Needle exchange programs {FAQs}”, Canadian Centre on substance abuse, http://www.ccsa.ca