Biostatistics (2015), **16**, 3, *pp.* 565–579
doi:10.1093/biostatistics/kxu061
Advance Access publication on January 18, 2015

Design-based inference in time-location sampling

LUCIE LEON*

French Institute for Public Health Surveillance, Saint-Maurice 94415, France
l.leon@invs.sante.fr

MARIE JAUFFRET-ROUSTIDE

French Institute for Public Health Surveillance, Saint-Maurice 94415, France and Cermes3, Inserm U988/UMR CNRS 8211/Ehess/Paris Descartes University, Paris

YANN LE STRAT

French Institute for Public Health Surveillance, Saint-Maurice 94415, France

Summary

Time-location sampling (TLS), also called time-space sampling or venue-based sampling is a sampling technique widely used in populations at high risk of infectious diseases. The principle is to reach individuals in places and at times where they gather. For example, men who have sex with men meet in gay venues at certain times of the day, and homeless people or drug users come together to take advantage of services provided to them (accommodation, care, meals). The statistical analysis of data coming from TLS surveys has been comprehensively discussed in the literature. Two issues of particular importance are the inclusion or not of sampling weights and how to deal with the frequency of venue attendance (FVA) of individuals during the course of the survey. The objective of this article is to present TLS in the context of sampling theory, to calculate sampling weights and to propose design-based inference taking into account the FVA. The properties of an estimator ignoring the FVA and of the design-based estimator are assessed and contrasted both through a simulation study and using real data from a recent cross-sectional survey conducted in France among drug users. We show that the estimators of prevalence or a total can be strongly biased if the FVA is ignored, while the design-based estimator taking FVA into account is unbiased even when declarative errors occur in the FVA.

Keywords: Hard-to-reach populations; Indirect sampling; Inference; Sampling weights; Time-location sampling; Venue-based sampling.

1. Introduction

Studying populations at high risk of infectious diseases is crucial to implement adequate prevention messages and interventions to reduce transmission. Drug users, men who have sex with men (MSM), sex workers, homeless people, and certain immigrants are examples of vulnerable populations particularly

*To whom correspondence should be addressed.

exposed to Hepatitis B and C, HIV, sexually transmitted infections, and other diseases. However, performing an epidemiological survey in these populations is difficult in many countries because of the illicit nature of certain practices, such as the use of drugs or prostitution. Specifically adapted sampling techniques have been developed over the past decades to survey such hard-to-reach populations (Sudman *and others*, 1988; Spreen, 1992; Thompson and Frank, 2000; Semaan *and others*, 2002; Magnani *and others*, 2005; Kalton, 1993; Tourangeau *and others*, 2014). One of these techniques, time-location sampling (TLS), also called time-space sampling or venue-based sampling, is widely used (Muhib *and others*, 2001; Stueve *and others*, 2001; Magnani *and others*, 2005), especially for surveys among MSM (Parsons *and others*, 2008; Paquette and De Wit, 2010; Paz-Bailey *and others*, 2014).

Pioneered in public health by the Centers for Disease Control and Prevention in the Young Men's Survey, TLS is a method for reaching individuals in places and at times where they congregate rather than where they live (MacKellar *and others*, 1996; Valleroy *and others*, 2000). Drug users are surveyed in specialized centers where they receive services (e.g. needle exchange, medical examinations, accommodation) (Jauffret-Roustide *and others*, 2009; Sutton *and others*, 2012). Homeless people are surveyed in support centers offering accommodation, care or free meals, or are surveyed in street locations (Chew *and others*, 2013). MSM are surveyed in gay venues (e.g. bars, clubs, saunas, etc.) (Wejnert *and others*, 2013).

Some authors have considered issues to ensure the validity of TLS in producing unbiased estimates in terms of proportions of individuals covered by surveys, the duration of the sampling period, the eligibility and the range (in terms of number of visits) of the venues, and the "representativeness" of the resulting sample (Stueve *and others*, 2001; Pollack *and others*, 2005; MacKellar *and others*, 2007; Parsons *and others*, 2008). The heterogeneity of the frequencies of venue attendance (FVA), also referred to as multiplicity, has often been highlighted and remains a major point of debate with respect to the efficacy and accuracy of TLS. Some individuals visit only one venue during the course of a survey while others visit dozens of venues in different places at different times. Literature has shown that among MSM these frequencies are heterogeneous from one individual to another, depending on several individual characteristics (Gustafson *and others*, 2013).

The first objective of this paper is to present TLS in the context of statistical sampling theory, which to our knowledge, has never yet been described. Although, some authors have introduced TLS as a "multi-step" procedure (Stueve *and others*, 2001; Pollack *and others*, 2005), it has only recently been presented as a two-stage or three-stage sampling design (Karon and Wejnert, 2012). Some authors still consider TLS a non-random sampling technique (Meyer and Wilson, 2009) and others have raised the question about whether it is necessary to weight or not to weight in TLS (Jenness *and others*, 2011; Xia and Torian, 2013; Risser and Montealegre, 2014). Our second objective is to investigate this latter point by introducing sampling weights which incorporate the FVA in a design-based estimator as an alternative to a recently proposed model-assisted estimator (Gustafson *and others*, 2013). Our estimator uses the indirect sampling framework and the generalized weight share method (GWSM) (Lavallée, 1995, 2007). The properties of an estimator ignoring the FVA and of the design-based estimator which takes FVA into account were assessed both by a simulation study and by using data from a national cross-sectional survey carried out in France among drug users in 2011. In addition, we explored the behavior of the alternative design-based estimator when errors occur in the FVA.

2. Time-location sampling

We focus on a population of individuals, named B, attending centers (locations) at certain times. We consider that these centers are open at various times during the survey period. For simplicity but without loss of generality, we consider that the opening time unit for the centers is a half-day. The following is also valid for other populations, irrespective of the type of center, the number of centers, and the time unit.

2.1 *Sampling design and sampling weights*

TLS can be viewed as a three-stage sampling design as illustrated in Figure 1. At the first stage, n centers are randomly drawn from a sampling frame of N centers indexed by l ($l = 1, \ldots, N$). At the second stage, for each center l ($l = 1, \ldots, n$) drawn at the first stage (named primary sampling unit (PSU)), we build a sampling frame of the N_l opening half-days during the survey period, indexed by k ($k = 1, \ldots, N_l$). We draw at random n_l half-days from the N_l half-days for each center l (named the secondary sampling units (SSUs)). We then establish a schedule representing each randomly drawn center and each randomly drawn half-day for the survey. To illustrate this, Figure 2 depicts an opening time schedule for 5 centers during a 4-week survey period. Finally, at the third stage, one or more investigators visit the centers according to the opening time schedule. For each center l and for each half-day k drawn, they randomly select n_{kl} among N_{kl} eligible individuals who attend these centers. Individuals represent the tertiary sampling units (TSUs).

At the first stage, either a simple random sampling without replacement (SRSWR) or an unequal random sampling without replacement is used. For the latter, the inclusion probability of a center is proportional to an available quantitative auxiliary variable, e.g. the average daily number of individuals attending the center. At the second and third stages, SRSWR is widely used. In most cases, the investigator does not have any list of individuals when arriving at a center. Systematic sampling is then often chosen as follows: the investigator randomly draws a person who arrives at the center, then selects the other individuals according to their ranking order of arrival using a sampling fraction defined *a priori*. Sometimes, a stratified sampling can also be employed, e.g. individuals are stratified by sex, age groups, nationalities, or any other characteristics of interest. The random selections of the sampling units at each stage (centers, half-days, individuals) aim to reduce selection biases.

To make inference in the population from the random sample, a sampling weight is assigned to each surveyed individual. The (first-order) inclusion probability for a unit is equal to the probability that this unit belongs to the sample. A sampling weight defined as the inverse of an inclusion probability can be expressed as the product of the sampling weights calculated at each stage of the design. We introduce the notation of the inclusion probabilities in Table 1 (column 2), under the assumption that an SRSWR is used at each stage. At the first stage, the sampling weight of a center l is $w_l = 1/\pi_l$. At the second stage, the sampling weight of a half-day k for the center l is $w_{k|l} = 1/\pi_{k|l}$. At the third stage, the sampling weight of an individual i surveyed in the center l during the half-day k is $w_{i|kl} = 1/\pi_{i|kl}$. The final inclusion probability of an individual i is $\pi_i = \pi_l \times \pi_{k|l} \times \pi_{i|kl}$ and his/her final sampling weight is:

$$w_i = w_l \times w_{k|l} \times w_{i|kl}. \tag{2.1}$$

2.2 *The Horvitz–Thompson estimator*

Very often, the main objective of cross-sectional surveys including time-location surveys is to estimate parameters of interest such as a total (e.g. population size, number of infected individuals), a proportion (e.g. proportion of infected individuals, called prevalence), or a mean (e.g. the mean value of a biomarker). For each individual i in the population B, let us consider a binary variable of interest y_i corresponding to his/her serological status for the disease of interest: y_i equals 1 if i is infected and 0 if not.

The Horvitz–Thompson estimator (Horvitz and Thompson, 1952) of the total number of infected individuals in the population $T = \sum_{i \in B} y_i$ is:

$$\hat{T} = \sum_{i \in s^B} w_i y_i, \tag{2.2}$$

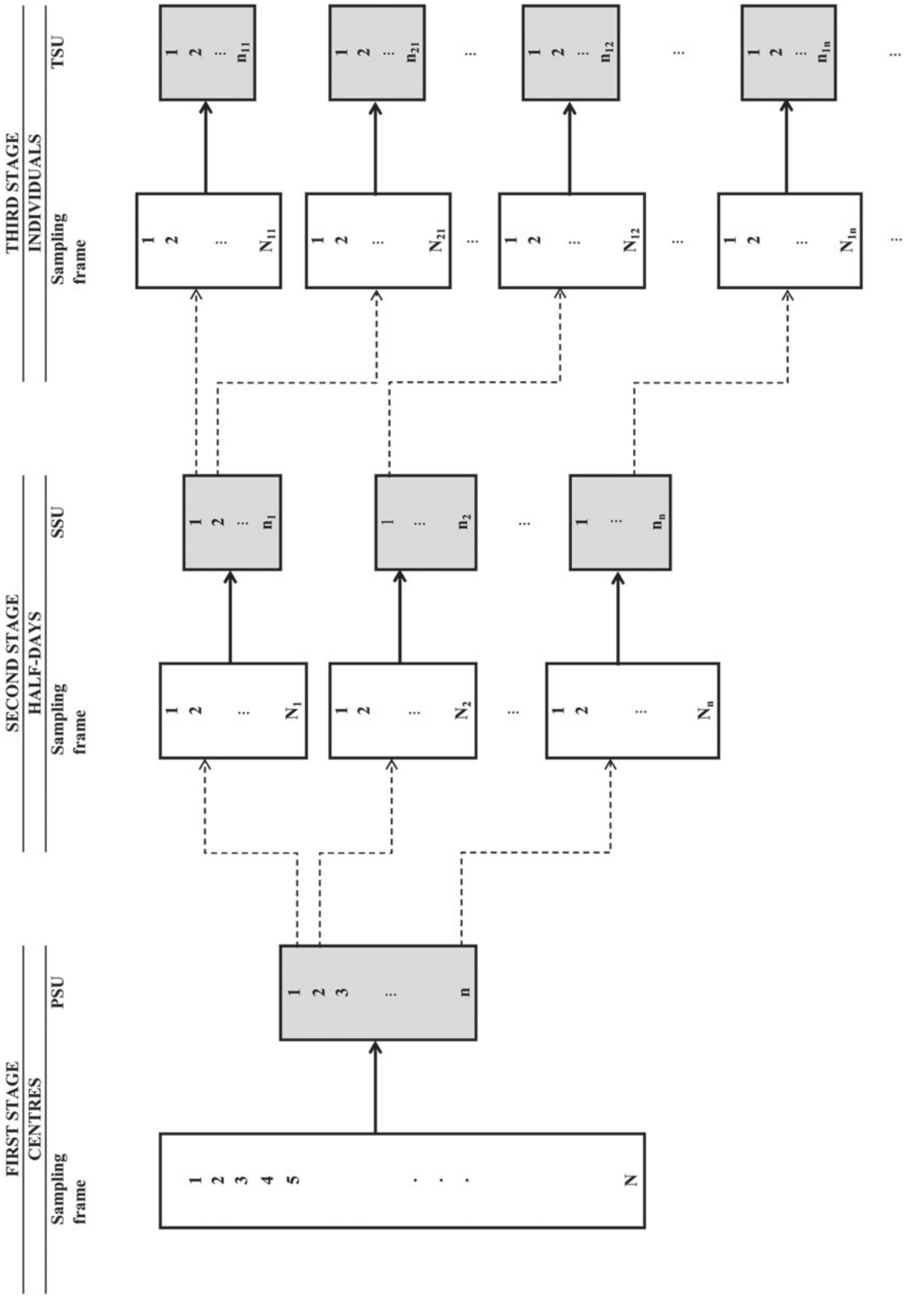
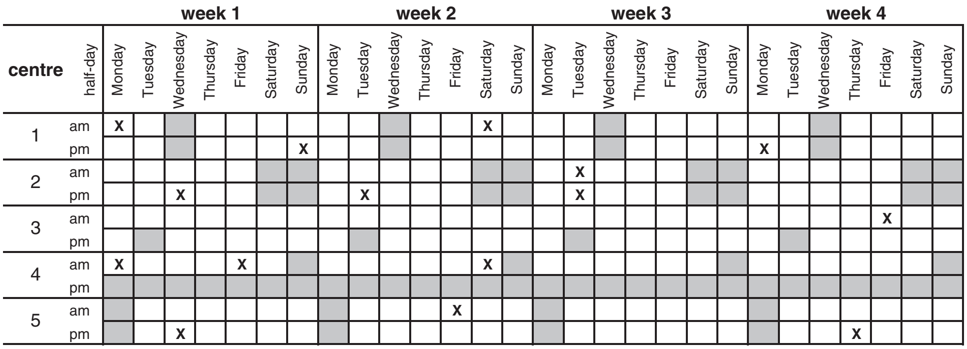


Fig. 1. Three-stage sampling design. Bold arrows represent the drawings and dashed arrows represent the sampling frames built for the units drawn at the first and second stages. PSU, primary sampling unit; SSU, secondary sampling unit; TSU, tertiary sampling unit.



centre	half-day	week 1							week 2							week 3							week 4						
		Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	am	X		▨							▨			X				▨							▨				
	pm			▨				X			▨							▨					X		▨				
2	am						▨	▨						▨	▨		X				▨	▨						▨	▨
	pm			X			▨	▨		X				▨	▨		X				▨	▨						▨	▨
3	am																										X		
	pm		▨							▨							▨							▨					
4	am	X				X		▨						X	▨							▨							▨
	pm	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨
5	am	▨							▨				X			▨							▨						
	pm	▨		X					▨							▨							▨			X			

Fig. 2. Schedule for 5 randomly drawn centers in a 4-week time-location survey. Centers are visited during the randomly drawn half-days (cross squares) among opening half-days (white squares). Gray squares represent the closing half-days.

Table 1. *Inclusion probabilities and totals expressions under the SRSWR assumption used at each stage of a three-stage sampling design*

Stage	First-order	Second-order[†]	Δ quantities	Totals
1	$\pi_l = \dfrac{n}{N}$	$\pi_{ll'} = \dfrac{n}{N}\left(\dfrac{n-1}{N-1}\right)$	$\Delta_{ll'} = \pi_{ll'} - \pi_l \pi_{l'}$	$T = \sum_{l=1}^{N} t_l$
2	$\pi_{k\vert l} = \dfrac{n_l}{N_l}$	$\pi_{kk'\vert l} = \dfrac{n_l}{N_l}\left(\dfrac{n_l-1}{N_l-1}\right)$	$\Delta_{kk'\vert l} = \pi_{kk'\vert l} - \pi_{k\vert l}\pi_{k'\vert l}$	$t_l = \sum_{k=1}^{N_l} t_{k\vert l}$
3	$\pi_{i\vert kl} = \dfrac{n_{kl}}{N_{kl}}$	$\pi_{ii'\vert kl} = \dfrac{n_{kl}}{N_{kl}}\left(\dfrac{n_{kl}-1}{N_{kl}-1}\right)$	$\Delta_{ii'\vert kl} = \pi_{ii'\vert kl} - \pi_{i\vert kl}\pi_{i'\vert kl}$	$t_{k\vert l} = \sum_{i=1}^{N_{kl}} y_i$

[†] $\pi_{ll} = \pi_l$; $\pi_{kk|l} = \pi_{k|l}$; $\pi_{ii|kl} = \pi_{i|kl}$.

where s^B is the sample drawn from the population B using TLS described above. The population size N^B, which is unknown in most cases, in particular for hard-to-reach individuals, is estimated by $\hat{N}^B = \sum_{i \in s^B} w_i$. The prevalence $P = T/N^B$ is estimated by:

$$\hat{P} = \frac{\hat{T}}{\hat{N}^B}. \tag{2.3}$$

The variance of a total estimator (Särndal *and others*, 2003), with respect to the sampling design, is estimated using the second-order inclusion probabilities (which constitute the joint inclusion probability of 2 distinct units) and other notations introduced to simplify the following formula (see Table 1, columns 3–5):

$$\widehat{Var}(\hat{T}) = \sum_{l=1}^{n}\sum_{l'=1}^{n} \Delta_{ll'} \frac{\hat{t}_l}{\pi_l}\frac{\hat{t}_{l'}}{\pi_{l'}} + \sum_{l=1}^{n} \frac{\widehat{Var}(\hat{t}_l)}{\pi_l} + \sum_{l=1}^{n} \frac{1}{\pi_l} \sum_{k=1}^{n_l} \frac{\widehat{Var}(\hat{t}_{k|l})}{\pi_{k|l}} \tag{2.4}$$

where l and l' denote 2 distinct centers, k and k' denote 2 distinct half-days, i and i' denote 2 distinct individuals and where

$$\hat{t}_l = \sum_{k=1}^{n_l} \frac{\hat{t}_{k|l}}{\pi_{k|l}}, \quad \hat{t}_{k|l} = \sum_{i=1}^{n_{kl}} \frac{y_i}{\pi_{i|kl}},$$

$$\widehat{\mathrm{Var}}(\hat{t}_l) = \sum_{k=1}^{n_l}\sum_{k'=1}^{n_l} \Delta_{kk'|l} \frac{\hat{t}_{k|l}}{\pi_{k|l}}\frac{\hat{t}_{k'|l}}{\pi_{k'|l}} \quad \text{and} \quad \widehat{\mathrm{Var}}(\hat{t}_{k|l}) = \sum_{i=1}^{n_{kl}}\sum_{i'=1}^{n_{kl}} \Delta_{ii'|kl} \frac{y_i}{\pi_{i|kl}}\frac{y_{i'}}{\pi_{i'|kl}}.$$

$\widehat{\mathrm{Var}}(\hat{N}^B)$ is calculated in a similar way by using (2.4) and assuming that $y_i = 1$ for any $i \in B$.

The estimated variance of the estimated prevalence is:

$$\widehat{\mathrm{Var}}(\hat{P}) = \widehat{\mathrm{Var}}\left(\frac{\hat{T}}{\hat{N}^B}\right) = \frac{1}{\hat{N}^{B^2}}\{\widehat{\mathrm{Var}}(\hat{T}) - 2\hat{P}\,\widehat{\mathrm{Cov}}(\hat{T}, \hat{N}^B) + \hat{P}^2\widehat{\mathrm{Var}}(\hat{N}^B)\}, \tag{2.5}$$

where

$$\widehat{\mathrm{Cov}}(\hat{T}, \hat{N}^B) = \sum_{l=1}^{n}\sum_{l'=1}^{n} \Delta_{ll'} \frac{\hat{t}_l}{\pi_l}\frac{\hat{N}_l}{\pi_{l'}} + \sum_{l=1}^{n} \frac{\widehat{\mathrm{Cov}}(\hat{t}_l, \hat{N}_l)}{\pi_l} + \sum_{l=1}^{n} \frac{1}{\pi_l} \sum_{k=1}^{n_l} \frac{\widehat{\mathrm{Cov}}(\hat{t}_{k|l}, \hat{N}_{k|l})}{\pi_{k|l}} \tag{2.6}$$

with

$$\hat{N}_l = \sum_{k=1}^{n_l} \frac{\hat{N}_{k|l}}{\pi_{k|l}}, \quad \hat{N}_{k|l} = \sum_{i=1}^{n_{kl}} \frac{1}{\pi_{i|kl}}, \quad \widehat{\mathrm{Cov}}(\hat{t}_l, \hat{N}_l) = \sum_{k=1}^{n_l}\sum_{k'=1}^{n_l} \Delta_{kk'|l} \frac{\hat{t}_{k|l}}{\pi_{k|l}}\frac{\hat{N}_{k'|l}}{\pi_{k'|l}}$$

and

$$\widehat{\mathrm{Cov}}(\hat{t}_{k|l}, \hat{N}_{k|l}) = \sum_{i=1}^{n_{kl}}\sum_{i'=1}^{n_{kl}} \Delta_{ii'|kl} \frac{y_i}{\pi_{i|kl}}\frac{1}{\pi_{i'|kl}}.$$

Note that if the second-order inclusion probabilities are easy to calculate when using SRSWR, their calculation is more complicated and sometimes intractable with other samplings and depend on the sampling algorithms used (Tillé, 2006). With more complex sampling designs, variances may be estimated using jackknife or bootstrap procedures (Särndal *and others*, 2003).

The Horvitz–Thompson estimator is unbiased for any sampling design if $\pi_i > 0$ for all $i \in B$ and of course if the inclusion probabilities are correctly calculated. For a population whose individuals attend several centers delivering services, the calculation of inclusion probabilities is more challenging than that for a population whose individuals are more static in time and space and who can be selected only once at most. In a time-location survey, the inclusion probability of an individual depends on his/her FVA.

In order to collect this information on FVA, we ask the respondents a set of questions to discover which venues they attend. One of the questions may be for example "how often did you go to any of the following venues during the previous 5 days?". Other more detailed questions may be asked according to the type of center (Gustafson *and others*, 2013). Then, the number of centers attended by each individual can be taken into account in a new estimator. As an alternative to the Horvitz–Thompson estimators ((2.2) and (2.3)) which can be biased when FVA is heterogeneous, we propose an unbiased design-based estimator incorporating the FVA. As this estimator is developed within the framework of indirect sampling, we firstly introduce the indirect sampling and then develop the new estimator.

3. Indirect sampling

Let us consider a population A containing N^A units indexed by j ($j = 1, \ldots, N^A$) and the population of interest B in which we want to estimate a function of interest (proportion, total) that contains N^B units indexed by i ($i = 1, \ldots, N^B$). A link between these 2 populations A and B is defined as the correspondence between any unit $j \in A$ with any unit $i \in B$ which allows switching back and forth between A and B. Indirect sampling designates a sampling for which: (1) a sample of units $j \in A$ named s^A is randomly drawn to access the units $i \in B$ and, (2) the units $i \in B$ are linked to, at least, one unit $j \in A$ (Deville and Lavallée, 2006; Lavallée, 2007). The correspondence between the 2 populations can be represented by a link matrix L of size $N^A \times N^B$. Each element $l_{ji} \geqslant 0$ defines the link between $i \in B$ and $j \in A$ and, if there is no link between the units, this quantity is 0. To illustrate indirect sampling, let consider a population A of 5 units and a population B of 3 units represented in Figure 1 in Section S.1 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org). The link matrix is:

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ 0 & l_{22} & 0 \\ l_{31} & 0 & 0 \\ 0 & 0 & l_{43} \\ 0 & 0 & l_{53} \end{pmatrix}.$$

A link l_{ji} is (1) bijective if a unit $i \in B$ has a one-to-one link with a unit $j \in A$, (2) injective if a unit $i \in B$ has at most one link with a unit $j \in A$, or (3) surjective if a unit $i \in B$ has at least one link with a unit $j \in A$.

Therefore, TLS can be viewed as a three-stage indirect sampling design where, at the third stage, population A is the population of services, population B is the population of individuals who receive these services in the centers and where the FVA of individuals is equal to the sum of the links between A and B, as illustrated in Section S.2 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org). The population of services offered by the centers exists but a list associated with this population is not available, except in special cases (e.g. accommodation). The estimator introduced in the following section is theoretically based on services received by the individuals surveyed. We will see in Section 5 how this estimator is calculated in practice even when we do not know which specific services individuals benefit from, and have only information about the centers they visited and how often.

In the framework of indirect sampling, we can use the GWSM, first described by Lavallée (1995), to provide a relevant sampling weight for each individual interviewed. A new sampling weight is assigned to each individual $i \in s^B$, basically defined as a weighted arithmetic mean of the sampling weights of the population of services involving the links between i and the services he/she received.

4. The alternative design-based estimator

The sampling weight of a service $j \in s^A$ is $w_j = w_i$ as defined in (2.1). If unit $j \in s^A$ is linked to unit $i \in s^B$, $l_{ji} \geqslant 0$ and if these 2 units are not related to each other, $l_{ji} = 0$. Note that some authors have highlighted the importance of the choices of link values that influence the precision of the estimates issued from indirect sampling, even though, in most applications, the values of l_{ji} for the linked units are equal to one (Lavallée and Caron, 2001; Deville and Lavallée, 2006). Then, for each unit $i \in s^B$, we can calculate the total number of links $L_i^B = \sum_{j \in A} l_{ji}$.

Finally, the final sampling weight incorporating the FVA for each unit $i \in s^B$ is defined as:

$$\tilde{w}_i = \frac{1}{L_i^B} \sum_{j \in s^A} l_{ji} w_i. \tag{4.1}$$

The alternative design-based estimators for the totals T and N^B and the prevalence P are, respectively:

$$\hat{T}_G = \sum_{i \in s^B} \tilde{w}_i y_i, \tag{4.2}$$

$$\hat{N}_G^B = \sum_{i \in s^B} \tilde{w}_i, \tag{4.3}$$

$$\hat{P}_G = \frac{\hat{T}_G}{\hat{N}_G^B}. \tag{4.4}$$

It has been demonstrated that these estimators are unbiased (Lavallée, 2007). Their respective variances are estimated using the same expressions proposed in (2.4)–(2.6) with

$$\widehat{\mathrm{Var}}(\hat{t}_{k|l}) = \sum_{j=1}^{n_{kl}} \sum_{j'=1}^{n_{kl}} \Delta_{ii'|kl} \frac{z_j}{\pi_{i|kl}} \frac{z_{j'}}{\pi_{i'|kl}},$$

$$\widehat{\mathrm{Cov}}(\hat{t}_{k|l}, \hat{N}_{k|l}) = \sum_{j=1}^{n_{kl}} \sum_{j'=1}^{n_{kl}} \Delta_{ii'|kl} \frac{z_j}{\pi_{i|kl}} \frac{1}{\pi_{i'|kl}}, \quad \text{and} \quad z_j = \sum_{i=1}^{n_i} \frac{l_{ji}}{L_i^B} y_i.$$

In real life, it is often pointless to ask participants what specific services they used or even their total number of visits over a survey period. First, individuals may be hesitant to answer this question. They go to centers for particular reasons and do not see why is it of any interest to spend time trying to remember what they did in the past, especially after a potentially long interview. This question can also be viewed by respondents as a check on illicit practices. Second, participants may find it difficult to answer such questions accurately due to forgetfulness or confusion as regards center identification. This is more marked in precarious populations or when there is a large number of centers (e.g. in a big city). Finally, the practical conditions of the interview rarely allow the collect of such detailed information, for example, when administering a questionnaire in the street or in a squat.

For all these reasons, researchers ask few questions about FVA over a short past period. This point is developed in greater detail in the next section.

However, it is important to note that we do not need detailed information about the services individuals benefit from to calculate the design-based estimator. Indeed, individuals are generally randomly drawn when they arrive at a center, irrespective of whether or not they are going to benefit from one or more services. Their inclusion probabilities therefore do not depend on the number of services they receive but on their number of visits to centers. Accordingly, we simply need to count the number of their visits at different centers.

Now, we will illustrate the properties of both the established Horvitz–Thompson and the alternative design-based estimators first using a cross-sectional survey (French ANRS-Coquelicot survey) conducted in 2011 and then using a comprehensive simulation study.

5. French ANRS-Coquelicot survey

5.1 Design

The French ANRS-Coquelicot survey was conducted in 2004 (Jauffret-Roustide *and others*, 2009) and in 2011 (Jauffret-Roustide *and others*, 2013) among drug users residing in metropolitan cities in France, to estimate the prevalence of hepatitis C virus (HCV) infection (based on serum testing), to assess the frequencies of at-risk practices and to follow the dynamics of the epidemic. In each city, we performed a comprehensive inventory of all centers providing services to drug users as follows: accommodation services including residential centers, hotel rooms, “sleep-in” centers (French social service accommodation centers), drug treatment centers including those providing methadone maintenance and psychotherapy, low threshold services including needle exchange programs and outreach work teams. We then constructed a sampling frame by each half-day that centers were open.

A two-stage TLS was used. All listed centers participated in the survey. At the first stage, we selected half-days in all centers using an SRSWR. At the second stage, at each center/half-day visit, drug users were selected using systematic random sampling (except for residential centers where all users were included in the survey). Participants were included if they provided written consent to be interviewed and to provide a self-obtained finger-prick blood samples in the form of a dried blood spot for HCV testing. Inclusion criteria for the survey were: >18 years of age, injected or snorted drugs “at least once during one’s life”, spoke French and agreed to participate in the survey by providing written, informed consent. The study questionnaire lasted approximately 45 min and was administered by professional interviewers with no ties to the recruitment centers. Interviewers had been trained for hard-to-reach populations and especially for interviews with drug users, in order to minimize social desirability bias associated with drug consumption and at-risk practices. We included 1568 drug users in the ANRS-Coquelicot study in 2011.

5.2 *Data collection regarding FVA*

As mentioned above, collecting an accurate list of services which each participant benefits from or the visits he/she makes over the whole survey period is unrealistic. Most researchers focus on asking few questions regarding FVA with some restrictions: the frequency of attendance is collected as a discrete variable (with some categories), over a short past period and sometimes using a limited number of centers (Karon and Wejnert, 2012; Gustafson *and others*, 2013).

In our survey, we asked 2 questions about FVA: (1) Yesterday and in the previous 3 days, did you attend one or more centers? If so, where and how many times? (2) Including the center where we are now, what other center or centers have you already attended today or do you intend on attending today? The FVA distribution in this survey is represented in Figure 4 in Section S.4.1 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org).

5.3 *Results*

In 2011, the survey was conducted over 11 weeks (May–July) in 121 centers and 1568 drug users were interviewed. The Horvitz–Thompson estimate of the population size was $\hat{T} = 48\,147$ (95% confidence interval [43 741; 52 553]) individuals while our design-based estimate was $\hat{T}_G = 43\,710$ (95% confidence interval [39 667; 47 753]). The Horvitz–Thompson estimate of the HCV prevalence among drug users was $\hat{P} = 43.4\%$, (95% confidence interval [39.3%; 47.6%]) while our design-based estimate was $\hat{P}_G = 43.7\%$, (95% confidence interval [39.5%; 47.9%]). The 2 estimates are close, probably because of the low variance of the number of links declared by drug users. This low observed variance may be due to measurement. We assumed that FVA did not vary over the 11-week survey period. We therefore decided to only collect data on FVA for the previous 5 days. Furthermore, drug users may be reluctant to answer to these questions

Table 2. *Parameters used to generate erroneous links*

Error	$L_i^{B,error}$	k
1	$L_i^B \times (k+1)$	$k \sim \mathcal{U}(-0.5, 0.5)$
2	$L_i^B + k$	$k \in [-(L_i^B - 1); L_i^B]$
3	$L_i^B \times (k+1)$	$k \sim \mathcal{U}(-0.5, 0)$

for the reasons described in Section 4. Errors in the declared FVA may occur, leading to a possible underestimation of variance.

To investigate the impact of these errors on the estimates in greater detail, we conducted a simulation that we present in the following section.

6. Simulation study

6.1 *Simulation process*

We generated several populations of individuals attending centers to benefit from one or more services during a fixed period. These simulated populations have prevalences ranging from 1% to 90% and the number of links (e.g. the FVA) depends or not on the serological status of each individual. Then, we generated 10 000 samples from each population. For each sample generated: $\hat{N}^B$, $\hat{T}$, $\hat{P}$, $\hat{N}_G^B$, $\hat{T}_G$, $\hat{P}_G$ were calculated. To explore the properties of our design-based estimator when errors occur in the FVA, we generated 3 kinds of errors, presented in Table 2. More details on the simulation process are given in Section S.3 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org).

6.2 *Results*

The simulation study shows that, for any scenario, the design-based estimator is unbiased irrespective of the prevalence (Figure 3, and Figures in Section S.6 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org)). On the contrary, the Horvitz–Thompson is biased for several scenarios, particularly for scenarios 13–16 where the FVA depends on serological status, as illustrated in Figure 3 for estimated prevalences and in Tables of Section S.5 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org) for estimated population sizes.

Figure 4 presents the relative bias for all the estimated prevalences. For scenarios 13–16, the estimated prevalences from the Horvitz–Thompson estimator, despite being unbiased such as those from our design-based estimator, are 1.05–2.22 times higher than the true prevalence.

Coverage probabilities of the estimated prevalences ranged from 87% to 100% using the alternative design-based estimator and from 0% (scenarios 13–16) to 95% using the Horvitz–Thompson estimator (values represented by circles in Section S.7 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org)).

When errors occurred in the declared FVA, we observed a low bias (scenarios 13–16) or sometimes an absence of bias (scenarios 1–12) in the estimations of prevalence using the alternative design-based estimator (Figures available in Section S.8 of supplementary material available at *Biostatistics* online (http://biostatistics.oxfordjournals.org)) irrespective of the true prevalence in the population. We expected that the observed bias would increase due to the kinds of errors introduced in the FVA and presented in Table 2 (link error 1 $\leqslant$ link error 3 $\leqslant$ link error 2) as illustrated in Section S.9 of supplementary material

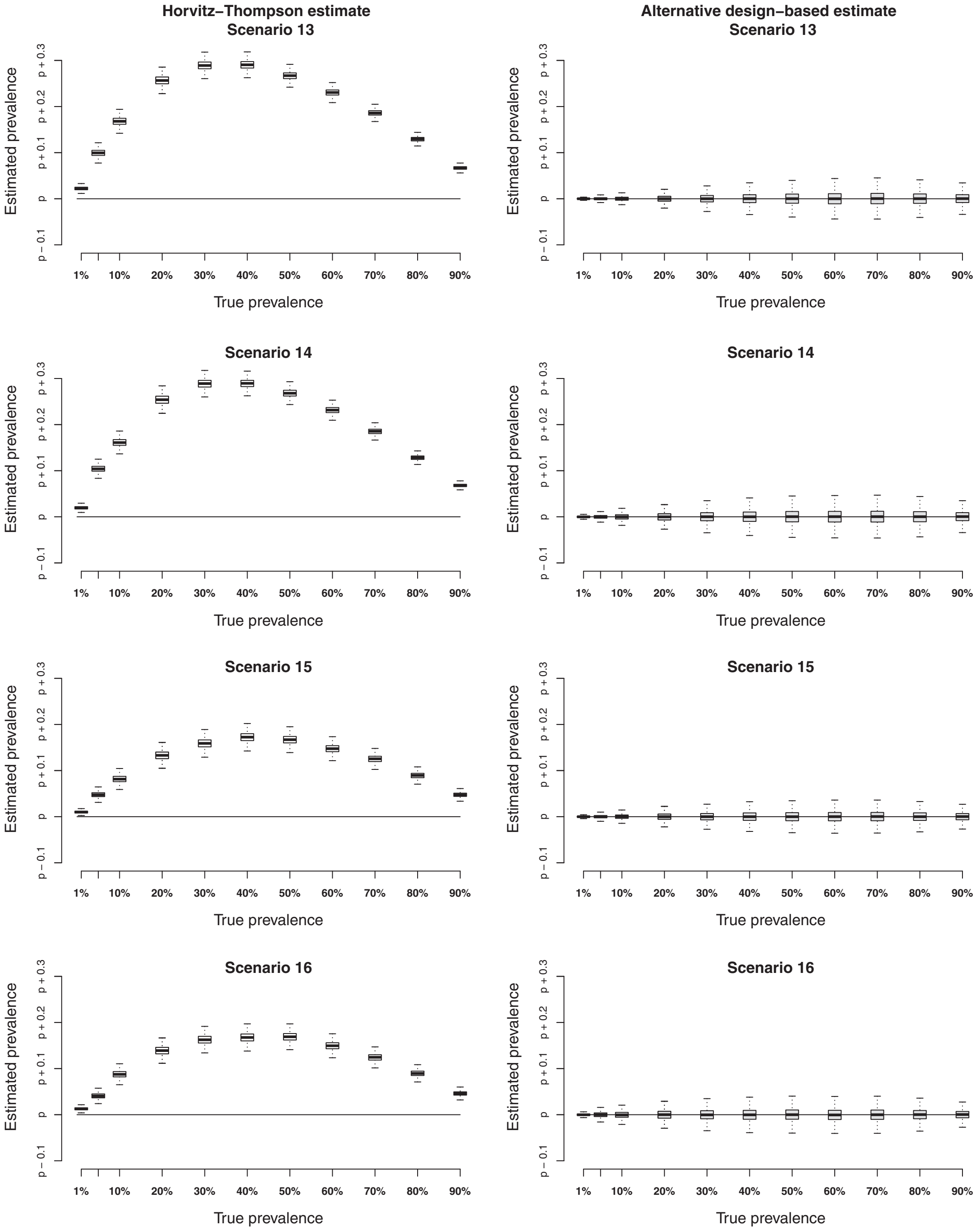


Fig. 3. Boxplots of estimated prevalences from the Horvitz–Thompson estimator (left) and from the alternative design-based estimator (right) for the scenarios 13–16. On each graph, the straight line represents the true prevalence in the simulated population for each scenario.

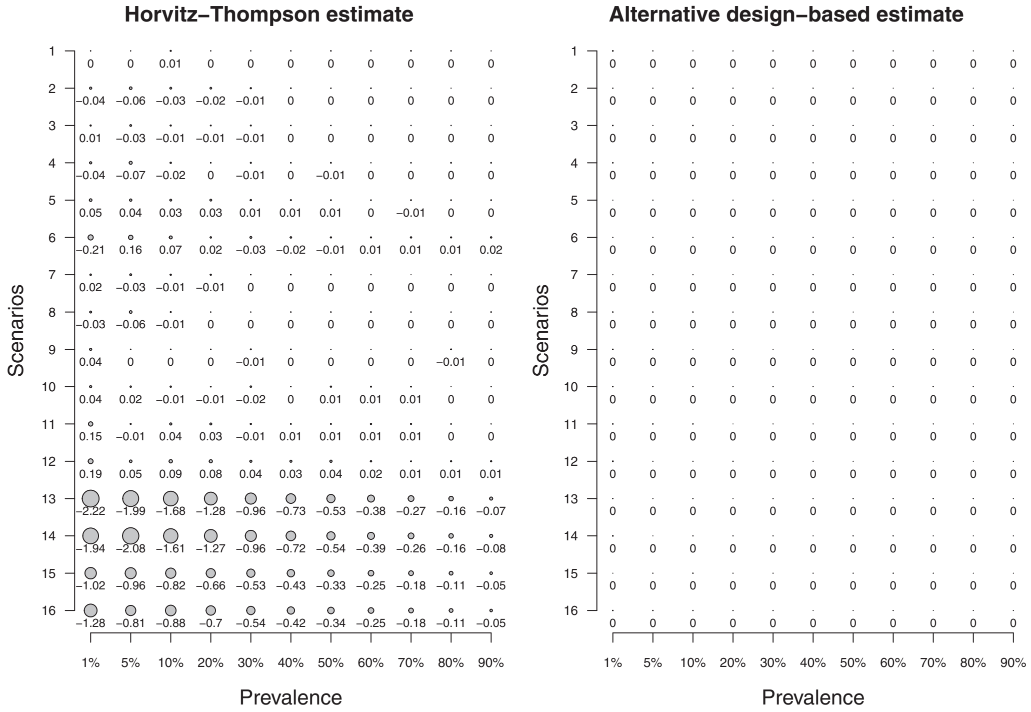


Fig. 4. Relative bias represented by the circles according to the scenarios and the different prevalences for the Horvitz–Thompson (left) and the alternative design-based (right) estimates.

available at *Biostatistics* online (http://biostatistics.oxfordjournals.org) when we estimated the number of infected individuals using both the Horvitz–Thompson and alternative design-based estimators.

7. DISCUSSION

We presented and implemented TLS as a multi-stage indirect sampling design and proposed a design-based estimator using the GWSM to provide accurate estimations for a total or a proportion when the population of interest in a survey is hard-to-reach and frequents specific venues. This design-based estimator takes into account the FVA of individuals, which was sometimes heterogeneous.

In the Coquelicot survey, the design-based estimator we proposed was adjusted for visits and showed results similar to those found using established Horvitz–Thompson estimator, due to the low variance in the FVA declared by participants. We did not carefully investigate why this observed variance was low but can put forward several explanations. First of all, the variance in the studied population of drug users was low. This is not the most likely explanation however as the participants had heterogeneous characteristics, particularly in terms of drug usage/consumption and therefore we expected them to have heterogeneous FVA. If this assumption is true, there is no benefit to using our estimator instead of the Horvitz–Thompson estimator in order to estimate proportions. However, the benefit is positive and real when estimating a total which must include the FVA.

A second most likely assumption is that the true variance is higher than that observed in the sample due to the difficulty in accurately collecting FVA. Indeed, participants with a great number of visits are not interested in spending time trying to recollect all their visits. The consequence is underestimated variance.

There is probably no perfect way to collect accurate information on FVA. It depends on the population studied and the surveyed locations. Future specific studies are needed to propose guideline questions to include related to FVA in questionnaires used in time-location surveys.

In the simulation study, we proposed different scenarios to cover several hard-to-reach populations with several prevalence values and with several FVA depending or not on the serological status. We concluded that collecting data on FVA during a face-to-face interview is crucial to modify the sampling weights in order to build an unbiased estimator. Even if errors occur in the FVA, the bias is reduced. Instead, ignoring FVA leads to severe bias and a weak coverage probability, in particular when FVA depends on serological status.

Our simulation mainly focused on the impact of FVA on the estimator bias. We did not investigate how other sources of bias could play a role in the robustness of the alternative design-based estimator. However, even if other sources of bias exist, our alternative estimator should always outperform the established Horvitz–Thompson estimator when FVA bias exists.

Furthermore, it could be interesting in a future extension of this study to compare our design-based estimator with the model-based estimator developed by Gustafson *and others* (2013) which focuses on simulated data, and to discuss the pros and the cons of these 2 estimators. From the results of the present study, we can already state that the use of indirect sampling coupled with the GWSM could solve several of the problems encountered in phone surveys when multiple communication with the same person because of both landline and mobile telephoning, must be taken into account.

8. Software

The ANRS-Coquelicot surveys were analyzed using STATA 12.1. The simulation study was implemented using the R software package (R version 3.0.2).

Supplementary material

Supplementary Material is available at http://biostatistics.oxfordjournals.org.

Acknowledgments

The authors thank the reviewer and the associate editors for their insight and helpful comments.
Conflict of Interest: None declared.

References

Chew, Ng. R. A., Muth, S. Q. and Auerswald, C. L. (2013). Impact of social network characteristics on shelter use among street youth in san francisco. *Journal of Adolescent Health* **53**, 381–386.

Deville, J. C. and Lavallée, P. (2006). Indirect sampling: the foundations of the Generalised Weight Share Method. *Survey Methodology* **32**, 165–176.

Gustafson, P., Gilbert, M., Xia, M., Michelow, W., Robert, W., Trussler, T., McGuire, M., Paquette, D., Moore, D. M. and Gustafson, R. (2013). Impact of statistical adjustment for frequency of venue attendance in a venue-based survey of men who have sex with men. *American Journal of Epidemiology* **177**(10), 1157–1164.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**, 663–685.

Jauffret-Roustide, M., Le Strat, Y., Couturier, E., Thierry, D., Rondy, M., Quaglia, M., Razafandratsima, N., Emmanuelli, J., Guibert, G., Barin, F. *and others*. (2009). A national cross-sectional study among drug-users in France: epidemiology of HCV and highlight on practical and statistical aspects of the design. *BMC Infectious Diseases* **9**, 113.

Jauffret-Roustide, M., Pillonel, J., Weill-Barillet, L., Léon, L., Le Strat, Y., Brunet, S., Benoit, T., Chauvin, C., Lebreton, M., Barin, F. *and others*. (2013). Estimation de la séroprévalence du VIH et de l'hépatite C chez les usagers de drogues en France - premiers résultats de l'enquête ANRS-COQUELICOT 2011. *Bulletin Epidémiologique Hebdomadaire* **39–40**, 504–509.

Jenness, S. M., Neaigus, A., Murrill, C. S., Gelpi-Acosta, C., Wendel, T. and Hagan, H. (2011). Recruitment-adjusted estimates of HIV prevalence and risk among men who have sex with men: effects of weighting venue-based sampling data. *Public Health Reports* **126**, 635–642.

Kalton, G. (1993). Sampling considerations in research on HIV risk and illness. *Methodological Issues in AIDS Behavioral Research*. New York: Plenum Press.

Karon, J. M. and Wejnert, C. (2012). Statistical methods for the analysis of time-location sampling data. *Journal of Urban Health* **89**, 565–586.

Lavallée, P. (1995). Cross-sectional weighting of longitudinal surveys of individuals and households using the weight share method. *Survey Methodology* **21**, 25–32.

Lavallée, P. (2007). *Indirect Sampling*. New York: Springer.

Lavallée, P and Caron, P. (2001). Estimation using the Generalised Weight Share Method: the case of record linkage. *Survey Methodology* **27**, 155–169.

MacKellar, D. A., Gallagher, K. M., Finlayson, T., Sanchez, T., Lansky, A. and Sullivan, P. S. (2007). Surveillance of HIV risk and prevention behaviors of men who have sex with men—a national application of venue-based, time-space sampling. *Public Health Reports* **122**(Suppl. 1), 39–47.

MacKellar, D., Valleroy, L., Karon, J., Lemp, G. and Jansen, R. (1996). The young men's survey: methods for estimating HIV seroprevalence and risk factors among young men who have sex with men. *Public Health Reports* **111**, 138–144.

Magnani, R., Sabin, K., Saidel, T. and Heckathorn, D. (2005). Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* **19** (Suppl. 2), S67–S72.

Meyer, I. H. and Wilson, P. A. (2009). Sampling lesbian, gay, and bisexual populations. *Journal of Counseling Psychology* **56**, 23–31.

Muhib, F. B., Lin, L. S., Stueve, A., Miller, R. L., Ford, W. L., Johnson, W. D. and Smith, P. J. (2001). A venue-based method for sampling hard-to-reach populations. *Public Health Reports* **116** (Suppl. 1), 216–222.

Paquette, D. and De Wit, J. (2010). Sampling methods used in developed countries for behavioural surveillance among men who have sex with men. *AIDS and Behavior* **14**, 1252–1264.

Parsons, J. T., Grov, C. and Kelly, B. C. (2008). Comparing the effectiveness of two forms of time-space sampling to identify club drug-using young adults. *Journal of Drug Issues* **38**, 1061–1082.

Paz-Bailey, G., Pham, H., Oster, A. M., Lansky, A., Bingham, T., Wiegand, R. E., Dinenno, E., Skarbinski, J. and Heffelfinger, J. D. (2014). Engagement in HIV care among HIV-positive men who have sex with men from 21 cities in the United States. *AIDS and Behavior* **18** (Suppl. 3), 348–358.

Pollack, L. M., Osmond, D. H., Paul, J. P. and Catania, J. A. (2005). Evaluation of the center for disease control and prevention's HIV behavioral surveillance of men who have sex with men: sampling issues. *Sexually Transmitted Diseases* **32**, 581–589.

Risser, J. M. and Montealegre, J. R. (2014). Comparison of surveillance sample demographics over two cycles of the National HIV Behavioral Surveillance Project, houston, texas. *AIDS and Behavior* **18** (Suppl. 3), 382–390.

Särndal, C. E., Swensson, B. and J. Wretman, J. (2003). *Model Assisted Survey Sampling*. New York: Springer.

Semaan, S., Lauby, J. and Leibman, J. (2002). Street and network sampling in evaluation studies of HIV risk-reduction interventions. *AIDS Reviews* **4**, 213–223.

Spreen, M. (1992). Rare populations, hidden populations, and link-tracing designs: what and why? *Sociological Methodology* **36**(1), 34–58.

Stueve, A., O'Donnell, L. N., Duran, R., Doval, A. S. and Blome, J. (2001). Time-space sampling in minority communities: results with young latino men who have sex with men. *American Journal of Public Health* **91**, 922–926.

Sudman, S., Sirken, M. G. and Cowan, C. D. (1988). Sampling rare and elusive populations. *Science* **240**, 991–996.

Sutton, A. J., McDonald, S. A., Palmateer, N., Taylor, A. and Hutchinson, S. J. (2012). Estimating the variability in the risk of infection for hepatitis C in the Glasgow injecting drug user population. *Epidemiology and Infection* **140**, 2190–2198.

Thompson, S. K. and Frank, O. (2000). Model-based estimation with link-tracing sampling designs. *Survey Methodology* **26**, 87–98.

Tillé, Y. (2006). *Sampling Algorithms*. New York: Springer.

Tourangeau, R., Edwards, B., Johnson, T. P., Wolter, K. M. and Bates, N. (2014). *Hard-to-Survey Populations*. Cambridge: Cambridge University Press.

Valleroy, L. A., MacKellar, D. A., Karon, J. M., Rosen, D. H., McFarland, W., Shedan, D. A., Stoyanoff, S. R., LaLota, M., Celentano, D. D., Koblin, B. A. *and others*. (2000). HIV prevalence and associated risks in young men who have sex with men. *Journal of the American Medical Association* **284**, 198–204.

Wejnert, C., Le, B., Rose, C. E., Oster, A. M., Smith, A. J., Zhu, J. and Paz-Bailey, G. for the NHBS study group. (2013). HIV infection and awareness among men who have sex with men-20 cities, United States, 2008 and 2011. *Plos One* **8**(10), 1–9.

Xia, Q. and Torian, L. V. (2013). To weight or not to weight in time-location sampling: why not do both? *AIDS and Behavior* **17**, 3120–3123.

[*Received July 8, 2014; revised December 15, 2014; accepted for publication December 17, 2014*]