Towards the Estimation of Effect Measures in Studies Using Respondent-Driven Sampling

Michael A. Rotondi

ABSTRACT *Respondent-driven sampling (RDS) is an increasingly common sampling technique to recruit hidden populations. Statistical methods for RDS are not straightforward due to the correlation between individual outcomes and subject weighting; thus, analyses are typically limited to estimation of population proportions. This manuscript applies the method of variance estimates recovery (MOVER) to construct confidence intervals for effect measures such as risk difference (difference of proportions) or relative risk in studies using RDS. To illustrate the approach, MOVER is used to construct confidence intervals for differences in the prevalence of demographic characteristics between an RDS study and convenience study of injection drug users. MOVER is then applied to obtain a confidence interval for the relative risk between education levels and HIV seropositivity and current infection with syphilis, respectively. This approach provides a simple method to construct confidence intervals for effect measures in RDS studies. Since it only relies on a proportion and appropriate confidence limits, it can also be applied to previously published manuscripts.*

INTRODUCTION

Respondent-driven sampling (RDS)^{1,2} is a chain-referral sampling strategy designed to recruit hidden populations. In RDS studies, recruitment begins with the selection of seeds who are selected from the target population of interest; each seed then recruits up to k members of the population, who in turn recruit up to k participants and so forth. Using Markov chain theory and the derivation of appropriate sampling weights, Heckathorn¹ showed that sample proportions (e.g., prevalence of a specific trait) in the sample will reach an equilibrium whereby they are no longer influenced by the initial choice of seeds. This allows for the creation of asymptotically unbiased estimates of the population proportion of interest, while a bootstrap approach provides appropriate confidence intervals (CI). Given its reliance on peer recruitment, this technique has facilitated sampling and study of numerous hard-to-reach populations, including HIV-positive individuals,³ transgender individuals,⁴ injection drug users,⁵ and jazz musicians.⁶ The most crucial assumption underlying successful use of RDS is that the population of interest be socially connected, as recruitment into the study must arise from one's peers.

Rotondi is with the York University, Toronto, Ontario, Canada.

Correspondence: Michael Rotondi, York University, Toronto, Ontario, Canada. (E-mail: mrotondi@yorku.ca)

A notable limitation of this analysis strategy is that the current approach and accompanying software⁷ provide only confidence intervals for the proportion of interest and do not readily provide point estimates or appropriate confidence intervals for effect measures, such as risk difference or relative risk. While the creation of a point estimate of the relevant effect measure is straightforward, construction of an appropriate CI is significantly more challenging. This paper thus describes an appropriate remedy to the CI construction problem based on the method of variance estimates recovery (MOVER).⁸ Briefly, this technique recovers the variance estimates from the upper and lower limits separately allowing for asymmetric confidence interval construction. This is of significant advantage in the analysis of RDS data as the bootstrap CIs obtained from analysis software (e.g., RDSAT⁷) are typically asymmetric. This approach has been applied to a variety of confidence interval estimation problems, including estimation of confidence intervals for functions of normal means and standard deviations,⁹ effect measures,¹⁰ and more generally, linear combinations of parameters.¹¹

Moreover, the construction of effect measures and confidence intervals using MOVER better quantifies the difference in risk between groups and has higher power than the overlap method, which is known to have low power to conclude that two groups in fact differ.¹² This is a particular strength when working with RDS data as this study design typically has larger variances than studies using simple random sampling,¹³ and thus benefits from this increase in power. The goal of this paper is twofold: (1) to provide a viable approach to constructing confidence intervals for the risk difference and relative risk using MOVER and (2) to illustrate this technique in the context of two examples, which are obtained from studies using RDS.

METHODS

The MOVER¹⁰ approach relies on recovering an estimate of the variance of each proportion near its confidence limits. This information is then used to construct a $100(1-\alpha)\%$ CI for the effect measure of interest, typically the risk difference or relative risk. Of particular strength is that it does not rely on any distributional assumptions on how the confidence limits for each parameter are obtained and thus may be used in the RDS context.¹⁰

Let $\hat{\pi}_1$ denote the point estimate of the first population proportion of interest, π_1 , with $100(1-\alpha)\%$ confidence limits, l_1 and u_1 , respectively. In a similar spirit, let $\hat{\pi}_2$ denote the point estimate of the second population proportion of interest, π_2 , with corresponding $100(1-\alpha)\%$ confidence limits, l_2 and u_2 . Assuming the nominal type I error rate of $\alpha=0.05$, this corresponds to the typical 95 % CI in practice. Note that depending on the analysis context, the populations may be different subgroups for which the investigator would like to compare the prevalence of a specified outcome or the prevalence of a trait across independent studies. Although the variances of $\hat{\pi}_1$ and $\hat{\pi}_2$ are not typically provided by RDS software (e.g., RDSAT⁷), available analysis software typically provide confidence intervals for each proportion of interest at the desired significance level using a bootstrap-resampling procedure.

Assuming $\hat{\pi}_1$ and $\hat{\pi}_2$ are independent, Zou and Donner¹⁰ provide us with a simple computational form for the $100(1-\alpha)\%$ confidence limits, L and U , for the risk difference, $\pi_1 - \pi_2$, where

$$L = \hat{\pi}_1 - \hat{\pi}_2 - \sqrt{\left(\hat{\pi}_1 - l_1\right)^2 + \left(u_2 - \hat{\pi}_2\right)^2} \quad (1)$$

and

$$U = \hat{\pi}_1 - \hat{\pi}_2 + \sqrt{\left(u_1 - \hat{\pi}_1\right)^2 + \left(\hat{\pi}_2 - l_2\right)^2} \quad (2)$$

In a similar spirit, the $100(1-\alpha)\%$ confidence limits can be obtained for the logarithm of risk ratio, $\ln(\pi_1/\pi_2)$ as

$$l = \ln(\hat{\pi}_1) - \ln(\hat{\pi}_2) - \sqrt{\left[\ln(\hat{\pi}_1) - \ln(l_1)\right]^2 + \left[\ln(u_2) - \ln(\hat{\pi}_2)\right]^2} \quad (3)$$

and

$$u = \ln(\hat{\pi}_1) - \ln(\hat{\pi}_2) + \sqrt{\left[\ln(u_1) - \ln(\hat{\pi}_1)\right]^2 + \left[\ln \pi_2 - \ln(l_2)\right]^2} \quad (4)$$

The required confidence limits for the risk ratio may now be simply obtained by $(\exp(l), \exp(u))$, as required.

EXAMPLES

The MOVER approach was applied to two recent RDS studies to generate confidence intervals for an appropriate effect measure of interest. In the first example, Kral et al.¹⁴ compare demographic characteristics of two samples of injection drug users (IDUs) in the San Francisco area. In our second example, Zhong et al.¹⁵ use RDS to study men who have sex with men (MSM) in Guangzhou, China, in order to determine risk factors that may be associated with HIV seropositivity and current infection with syphilis.

IDUs in San Francisco, USA

Kral et al.¹⁴ compare the results of two studies of IDUs, the first using targeted sampling and the second RDS, in the San Francisco area. The purpose of this report was to determine whether the studies recruited a representative sample of the IDU population. To assess this, comparisons were made on three general domains of interest: (1) sociodemographic characteristics, (2) reported drug of choice, and (3) access to and use of HIV prevention and treatment programs, including drug and alcohol treatment programs. For reference, the RDS study had a total sample size of 534 participants, while the convenience study recruited 651 individuals. The comparisons of both the RDS and targeted sample are presented in Table 1 of Kral et al.¹⁴ The authors used the overlap method to examine differences between the traits. That is, the authors compared the 95 % CIs between the same variables in the two studies to determine overlap between the CIs and concluded that those that did not overlap were statistically significantly different at the nominal 5 % type I error rate. Comparisons that were shown to be statistically significantly different in Kral et al.¹⁴ are italicized in Table 1. Additional differences found using MOVER are highlighted in boldface.

To illustrate the proposed technique to evaluate differences in proportions, we consider three variables of interest: (1) race/ethnicity, (2) ever tested positive for hepatitis C (HCV), and (3) ever participated in alcohol/drug treatment programs.

TABLE 1 Comparison of demographic characteristics between RDS and convenience samples from Kral et al.¹⁴

	RDS study	Convenience sample	MOVER approach
	Adjusted, % (95 % CI)	Unadjusted, % (95 % CI)	Difference in proportions, % (95 % CI)
Race/ethnicity			
Native American	5.8 (2.2, 9.5)	2.5 (1.5, 4.0)	3 (−1, 7)
Asian	0.6 (0.0, 1.9)	0.2 (0.0, 1.0)	0 (−1, 2)
African American	35.7 (29.3, 44.0)	49.3 (45.0, 53.0)	−14 (−21, −4)
Native Hawaiian/Pacific	0.5 (0.0, 1.2)	0.3 (0.0, 1.2)	0 (−1, 1)
Non-Hispanic White	35.3 (28.9, 41.8)	35.8 (32.0 ^a , 39.6)	−1 (−8, 7)
Hispanic	12.9 (6.7, 18.6)	7.3 (5.4, 9.6)	6 (−1, 12)
Other	1.5 (0.5, 3.3)	4.8 (3.3, 6.8)	−3 (−6, −1)
Ever tested positive for HCV	55.9 (49.8, 63.1)	69.7 (65.9, 73.1)	−14 (−21, −6)
Ever participated in alcohol/drug treatment programs	65.7 (60.3, 71.9)	74.8 (71.3, 78.1)	−9 (−15, −2)

Statistically significant differences found using the overlap method are italicized. Additional differences found using MOVER are highlighted in boldface

^aKral et al.¹⁴ stated 3.2 %, which should read 32.0 %

The results of these calculations (Eqs. 1 and 2) are summarized in Table 1. In addition to quantifying the differences in proportions, the MOVER approach also found a statistically significant difference in the proportion of “Other” races and the proportion of individuals who have “Ever participated in alcohol/drug treatment programs” in the two samples. These results are of interest as the original paper could not determine statistical significance for these outcomes using the overlap method due to low power.¹² Note that MOVER results are rounded to the nearest percent to avoid exaggerating precision.¹⁶ This is particularly relevant in RDS studies as the default radius of convergence is 2 %.⁷

HIV in MSM in Guangzhou, China

Zhong et al.¹⁵ used RDS to recruit 379 MSM in Guangzhou, China, to determine both the prevalence of and risk factors for HIV seropositivity and current syphilis infections. The authors estimated the prevalence of HIV and syphilis infection outcomes, and bivariate associations between HIV seropositivity and its risk factors, as well as current syphilis infection and its risk factors. Data from this published report were used to illustrate the estimation of the relative risk and CI for education levels and their association with HIV status and syphilis infection.

Applying the formulae for the 95 % CI for the log relative risk (Eqs. 3 and 4), we can estimate that individuals who had college or higher education status were 3.42 times less likely to be infected with syphilis (Table 2). While there is no evidence of a statistically significant difference between education levels and HIV seropositivity, this approach better quantifies the association through its inclusion of point estimates for the relative risk and CIs instead of simple *P* values. In this way, MOVER also provides an estimate of the *magnitude* of the effect size in addition to the ascertainment of statistical significance.

TABLE 2 Estimation of relative risks for HIV and syphilis infection for education levels (Zhong et al.¹⁵)

	HIV seropositivity		Current syphilis infection	
	Adjusted, % (95 % CI)	Relative risk (95 % CI)	Adjusted, % (95 % CI)	Relative risk (95 % CI)
Education level				
Junior high school or lower	7.7 (1.9, 14.7)	Referent	28.7 (19.2, 37.8)	Referent
Senior high school, technical	6.1 (1.9, 11.8)	1.3 (0.3, 4.8)	26.2 (17.4, 34.3)	1.1 (0.7, 1.8)
College or higher	2.7 (0.0, 6.2)	2.8 (0.6, 82)	8.4 (5.3, 13.7)	3.4 (1.8, 5.8)

DISCUSSION

Although methodological development is ongoing in multiple areas of RDS, including sample size estimation^{13,17} and understanding the potential for bias in RDS studies,^{18,19} respondent-driven sampling remains an appropriate recruitment strategy for hidden, networked populations. This manuscript provides a simple technique to obtain appropriate confidence limits for an effect measure of interest, namely the risk difference (difference in proportions) or relative risk. This is of significant value given the merits of presenting confidence intervals over simple hypothesis tests (e.g., Gardner and Altman²⁰) and the importance of estimating the strength of associations in epidemiological studies (e.g., Rothman and Greenland,²¹ Ch. 4). Moreover, this approach can easily add to the sophistication of RDS analyses without undue technical challenges.

The majority of RDS studies are cross-sectional in nature; thus, it may be desirable to obtain estimates of the difference in proportions to compare prevalence over multiple geographic regions or across multiple subgroups. This problem may now easily be examined using the MOVER approach to estimate differences between proportions and their associated CIs. In other RDS analyses, where an investigator is interested in quantifying the association between a potential outcome and risk factor of interest, the relative risk may be an appropriate effect measure.

A particular strength of this approach is that it eliminates the need to use ad hoc techniques such as comparing the overlap of confidence intervals to establish statistical significance in the comparison of two groups. This is particularly useful given the known limitations of this approach, including a lack of power to conclude that the groups differ and the conservative type I error rate when group differences are in fact found to differ.¹² Finally, the required data (point estimates and CIs) are typically published in manuscripts, and these straightforward computations can be performed using either a simple spreadsheet or statistical software.

REFERENCES

1. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc Probl.* 1997; 44: 174–199.

2. Heckathorn DD. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Soc Probl.* 2002; 49: 11–34.
3. Ramirez-Valles J, Garcia D, Campbell RT, et al. HIV infection, sexual risk behavior, and substance use among latino gay and bisexual men and transgender persons. *Am J Public Health.* 2008; 98: 1036–1042.
4. Rotondi NK, Bauer GR, Scanlon K, et al. Prevalence of and risk and protective factors for depression in female-to-male transgender Ontarians: trans PULSE Project. *Can J Commun Mental Health.* 2011; 30: 135–155.
5. Heckathorn DD, Semaan S, Broadhead RS, et al. Extensions of respondent-driven sampling: a new approach to the study of injection drug users aged 18–25. *AIDS Behav.* 2002; 6: 55–67.
6. Heckathorn DD, Jeffri J. Finding the beat: using respondent-driven sampling to study jazz musicians. *Poetics.* 2001; 28: 307–329.
7. Volz E, Wejnert C, Cameron C, et al. *Respondent-driven sampling analysis tool (RDSAT) version 7.1.* Ithaca: Cornell University; 2012.
8. Zou GY. On the estimation of additive interaction by use of the four-by-two table and beyond. *Am J Epidemiol.* 2008; 168: 212–224.
9. Donner A, Zou GY. Closed-form confidence intervals for functions of the normal mean and standard deviation. *Stat Methods Med Res.* 2012; 21: 347–359.
10. Zou GY, Donner A. Construction of confidence limits about effect measures: a general approach. *Stat Med.* 2008; 27: 1693–1702.
11. Newcombe RG. Propagating imprecision: combining confidence intervals from independent sources. *Commun Stat A-Theor.* 2011; 40: 3154–3180.
12. Schenker N, Gentleman JF. On judging the significance of differences by examining the overlap between confidence intervals. *Am Stat.* 2001; 55: 182–186.
13. Salganik MJ. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *J Urban Health.* 2006; 83: 98–112.
14. Kral AH, Malekinejad M, Vaudrey J, et al. Comparing respondent-driven sampling and targeted sampling methods of recruiting injection drug users in San Francisco. *J Urban Health.* 2010; 87: 839–850.
15. Zhong F, Lin P, Xu H, et al. Possible increase in HIV and syphilis prevalence among men who have sex with men in Guangzhou, China: results from a respondent-driven sampling survey. *AIDS Behav.* 2011; 15: 1058–1066.
16. Wilcox AJ. On precision. *Epidemiology.* 2004; 15: 1.
17. Wejnert C, Pham H, Krishna N, et al. Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States. *AIDS Behav.* 2012; 16: 797–806.
18. McCreesh N, Frost SD, Seeley J, et al. Evaluation of respondent-driven sampling. *Epidemiology.* 2012; 23: 138–147.
19. Goel S, Salganik MJ. Assessing respondent-driven sampling. *Proc Natl Acad Sci USA.* 2010; 107: 6743–6747.
20. Gardner MJ, Altman DG. Confidence intervals rather than p values: estimation rather than hypothesis testing. *BMJ.* 1986; 292: 746.
21. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology.* Philadelphia: Wolters Kluwer Health; 2008.