A response to the use of respondent-driven sampling in urban Indigenous populations

Dear Editor:

Re: Firestone M, Smylie J, Maracle S, McKnight C, Spiller M, O'Campo P. Mental health and substance use in an urban First Nations population in Hamilton, Ontario. *Can J Public Health* 2015;106(6):e375–e381. doi: 10.17269/cjph.106.4923

Critical review and evaluation of the article “Mental health and substance use in an urban First Nations population in Hamilton, Ontario” by Firestone et al.¹ brought an important issue to the forefront, namely the use of respondent-driven sampling for participant recruitment. The use of respondent-driven sampling during the study's recruitment period raises questions about the presence of any underlying biases that were not controlled for, nor addressed in any way throughout the article. This is problematic as unaddressed biases can hinder the external validity of the study.² This study in particular is at risk for developing sampling bias as a consequence of respondent-driven sampling.

Respondent-driven sampling is used when hard-to-reach populations are being studied.³ Although an accurate representation of the sample population is produced with respondent-driven sampling, the same cannot be said about its corresponding target population.³ Findings from respondent-driven sampling are often questionable with regard to their generalizability and reliability.⁴ Thus, caution must be considered when interpreting results, as respondent-driven sampling does not control for, or reduce, bias once it has occurred.³ Respondent-driven sampling was useful in reaching the hidden urban Aboriginal population in Hamilton, Ontario. It was also useful in collecting representative data on the urban Aboriginal sample population. However, problems with accuracy and representativeness arise when applying the study's results to other urban Aboriginal populations elsewhere in Canada.

Without abandoning respondent-driven sampling altogether, alternative considerations within its methodology could diminish the effects of sampling bias. The study used staff-selected seeds to participate and recruit additional participants.¹ This method is common in most respondent-driven sampling studies, but can also influence the further generalizability of studies.¹ The authors of the article could have selected an alternative method to recruit seeds and additional participants. That is, conducting respondent-driven sampling using self-presenting seeds and participant recruits. Self-presenting seeds and participant recruits approach the study independently (e.g., word-of-mouth) without explicit selection from researchers or staff (4). Unlike staff-selected seeds and participant recruits, there is evidence to support that self-selecting seeds reach a broader range and variation of the hard-to-reach population (1). While the self-presenting seeds and participants increase participant variation, they are able to enhance the accuracy and representativeness of respondent-driven sampling (4). Thus, the use of self-presenting seeds in the urban aboriginal population of Hamilton, Ontario could have expanded the range of

participants, making the results more generalizable to other aboriginal populations.

Jenna Osmond, BSc

Department of Community Health & Humanities, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL

Correspondence: Jenna Osmond, josmond@grenfell.mun.ca

doi: 10.17269/CJPH.108.6017

REFERENCES

1. Firestone M, Smylie J, Maracle S, McKnight C, Spiller M, O'Campo P. Mental health and substance use in an urban First Nations population in Hamilton, Ontario. *Can J Public Health* 2015;106(6):e375–e381. doi: 10.17269/cjph.106.4923.
2. Aschengrau A, Seage G. Bias. In Aschengrau A, Seage G (Eds.), *Essentials of Epidemiology in Public Health*. Burlington, MA: Jones & Bartlett Learning, 2014.
3. McCreesh N, Frost S, Seeley J, Katongole J, Tarsh MN, Ndungu R, et al. Evaluation of respondent-driven sampling. *Epidemiology* 2012;23(1):138–47. PMID: 22157309. doi: 10.1097/EDE.0b013e31823ac17c.
4. Wylie JL, Jolly AM. Understanding recruitment: Outcomes associated with alternate methods for seed selection in respondent driven sampling. *BMC Med Res Methodol* 2013;13(1):1–11. doi: 10.1186/1471-2288-13-93.

In response to Osmond letter

We thank Ms. Osmond for her feedback on our published CJPH article.¹ As outlined in our paper, Respondent Driven Sampling (RDS) was used to recruit 554 First Nations adults in Hamilton, Ontario. RDS draws on Markov chain theory, including the following theorems: 1) that in chain-referral processes, as recruitment continues from wave to wave, an equilibrium mix of recruits will be obtained that is independent of the characteristics of the seeds; 2) this occurs at a rapid rate, irrespective of the distribution of the initial seeds.² These theorems have been widely proven in hundreds of RDS studies to date.

Ms. Osmond expresses a concern regarding potential sampling bias and decreased generalizability of our results due to the use of staff-selected seeds. She proposes that self-presenting seeds would have increased participant variation and enhanced the representativeness of our study population. In fact, according to the proven Markov theorems described above, the characteristics of the seeds in RDS do not play a key role in the achievement of an equilibrium mix that is independent of the characteristics of these seeds. Self-presentation can also introduce a participant bias, as it is well known that people who self-select to volunteer for a study do not reflect the general population. Fortunately, in RDS, we do not have to worry about how seeds represent the study population, so self-presentation is a viable method.

Our community-based efforts to select seeds across various socio-economic strata and from diverse networks within the First Nations community were highly successful in our key objective in seeding, which was to engage diverse First Nations community members in the study. This proactive approach may have been more effective in community engagement than self-presentation, which is more passive.

In our sample, the two longest recruitment chains were comprised of 19 and 32 waves, respectively. Although newer tests such as bottle-neck plots could be used,³ this is a strong indication that we were able to overcome original sampling bias, which

usually requires only six or seven waves.⁴ Furthermore, anecdotal evidence from our community partner in Hamilton indicates that Indigenous community members who had not previously been accessing health services and programs and who were more “hidden” were now being recruited into the Our Health Counts study. We acknowledge that RDS has several assumptions that are difficult to assess³ and the limitations you describe regarding generalizability of our RDS study results are not unique to this study design and remain a challenge for all observational studies. We invite further dialogue about RDS as an evolving method and approach that builds on social networks and generates population-based health information.

Michelle Firestone, PhD, Janet Smylie, MD, MPH, FCFP

REFERENCES

1. Firestone M, Smylie J, Maracle S, McKnight C, Spiller M, O'Campo P. Mental health and substance use in an urban First Nations population in Hamilton, Ontario. *Can J Public Health* 2015;106(6):e375–81. PMID: 26680428. doi: 10.17269/cjph.106.4923.
2. Heckathorn D. Respondent-driven sampling: A new approach to the study of hidden populations. *Soc Probl* 1997;44:174–99. doi: 10.2307/3096941.
3. Gile KJ, Johnston LG, Salganik MJ. Diagnostics for respondent-driven sampling. *J R Stat Soc Ser A Stat Soc* 2015;178(1):249–69.
4. Wejnert C. An empirical test of respondent-driven sampling: Point estimates, variance, degree measures and out-of-equilibrium data. *Sociol Methodol* 2009;39(1):73–116. PMID: 20161130. doi: 10.1111/j.1467-9531.2009.01216.x.