Invited Commentary

Which is the cart and which is the horse? Getting more out of cross-sectional epidemiological studies

While several of the Bradford Hill criteria[1] for causation have been debated[2] since their description in 1965, few have disputed the fourth item, 'the temporal relationship of the association', where he posed the question: which is the cart and which is the horse? Temporality refers to the necessity for an exposure, or a hypothetical cause, to precede an outcome, or an effect, in time[1,3]. While such a criterion is inarguable in the field of causal inference[3], it is also important for permitting researchers to draw informative conclusions from their epidemiological analyses of associations[4]. For example, when devising interventions, it is often the goal to intervene upon an antecedent exposure to reduce the occurrence of an outcome[5]. If a researcher is generating evidence to help guide such decisions, it is important to attempt to quantify associations that are, at least in theory, representing a prospective sequence of events. Therefore, even in a cross-sectional study where there is often an inability to establish temporal ordering in the data[6], regression models typically generate more meaningful estimates if there is a hypothesized temporality of the relationships under study[3,4,7].

The paper by Patil *et al.*[8], published in this issue of *Public Health Nutrition*, is a cross-sectional study which examined food insecurity (FI) among people living with HIV (PLHIV) in the city of Pune in western India. Given the context-specific nature of FI[9], I read this paper with great interest and I commend the authors for their novel contribution to the literature. The objective of this paper was to assess the prevalence of FI along with risk factors for FI in a convenience clinical sample of 483 PLHIV (≥18 years of age) in 2015–2016. While cross-sectional studies are often described as being best suited for assessing prevalence[6], the authors also examined the relationships between several exposures and an outcome that were measured at a single point in time.

In our own work, we have documented the prevalence of FI and examined the relationship between injection drug use (IDU) and FI among people living with HIV–hepatitis C virus co-infection in Canada[10,11]. We considered past research which had suggested that the relationship between IDU and FI was bidirectional[12], whereby IDU may impact FI and FI may also be a risk factor for IDU. Therefore, to conduct an analysis of the relationship between these factors, with the goal of informing interventions to mitigate FI, it was important for us to hypothesize as to why the relationship may act in a given direction.

After an examination of the relevant literature[13], we decided to focus our attention on IDU as an exposure variable and FI as a potential consequence of this behaviour. We conceptualized this analysis[10], in time, using a directed acyclic graph, which is a visual representation of the hypothetical relationship between variables[7,14]. Unlike the cross-sectional study by Patil *et al.*[8], we benefited from the use longitudinal cohort data, which allowed us to temporally order the exposure (IDU) and outcome (FI) measures in our data set through variable lagging[10,11]. However, even in the absence of repeated measurements on participants, regression analyses typically generate more meaningful estimates if there is at least a hypothesized temporality of the relationships under examination[3,4,7]. In our papers, we concluded that a hypothetical intervention on IDU may potentially reduce FI[10,11]. While this conclusion was contingent upon several assumptions (e.g. that there were no unmeasured or imperfectly measured confounders), these studies also provided the foundation for future work in the same population. Specifically, we also found that methadone treatment, a substance use intervention, was associated with a lower risk of FI[15]. Such a study would not have been sufficiently motivated if we had not completed our earlier analyses with a temporal ordering of the IDU–FI relationship in mind.

In the study by Patil *et al.*[8], the following independent risk factors for FI were identified from a single adjusted logistic regression model and reported in the abstract: monthly family income and consuming ≥4 non-vegetarian meals per week. Estimates for other potential risk factors (i.e. age, sex, education and living location) were also reported in Table 1. In two separate logistic models that were adjusted for CD4 cell count, time on antiretroviral therapy, age, sex and HIV viral load, the authors indicated that two biomarkers, highly sensitive C-reactive protein (hs-CRP) and D-dimer, were independently associated with FI (Fig. 1); these findings were highlighted in the title of the paper and were emphasized in the concluding remarks. The choice to examine these two biomarkers, specifically, was seemingly driven by statistical significance in univariate analyses. Much like other cross-sectional studies, the authors correctly stated that their 'study is limited too in its failure to establish causality'. However, while this is an intrinsic limitation of their design[6], the use of cross-sectional data does not preclude researchers from motivating their analyses with a proposed temporal relationship between variables[4,7].

For example, while the finding related to inadequate income and FI was discussed with several references, there was little substantive rationale provided for studying the relationships between non-vegetarian meal consumption, hs-CRP and D-dimer as potential risk factors for FI. Regarding the meal consumption finding, it was stated that it was 'likely due to the higher costs and resources needed with obtaining and cooking non-vegetarian foods'[8]. While this may be true, it is unclear as to whether such a statement is grounded in existing evidence; dietary choices are typically described as a consequence of FI[16–19], as opposed to a risk factor or determinant of this experience.

Importantly, generating hypotheses regarding relationships between variables can be complicated by what is known as mutual adjustment[20] or the 'Table 2 Fallacy'[21]. Generally, the potential for misinterpretation and a lack of reproducibility are more common when multiple adjusted effect estimates are interpreted in a single regression model[20–23]. For example, in the author's Table 1[8] (which combines both a description of the study sample, a typical 'Table 1', and model estimates, a typical 'Table 2'), the adjusted odds ratios for all of the non-biomarker related factors are presented, where all variables were treated as potential FI risk factors. As such, the estimate for the 'consuming ≥4 non-vegetarian meals per week' variable was only one of several outputs presented and discussed from this multivariable model. This fallacy highlights that such exploratory modelling strategies may impede an author's ability to clearly motivate his/her analyses, to consider temporal ordering and confounding, and to contextualize the results[20–22]. As described, such issues can be addressed, in part, by using directed acyclic graphs to map out the potential relationship between an exposure and an outcome prior to regression modelling[7,14].

Regarding the hs-CRP and D-dimer findings presented in Fig. 1[8], it was discussed that 'it is possible that PLHIV with higher inflammation, and consequently higher levels of either biomarker, could have worse health-related quality of life[24], precluding them from achieving food security'. While there is an implicit temporality to this statement (i.e. biomarkers as a proxy for lower health-related quality of life leading to FI), the authors did not reference literature which lends support to such a pathway. In fact, some existing evidence hypothesizes and demonstrates that lower health-related quality of life is more likely a consequence[25–28], rather than a determinant, of FI. The hs-CRP and D-dimer findings are further complicated by the FI recall period referring to the four weeks prior to the administration of the Household Food Insecurity Access Scale; this scale was administered at the time of enrolment[8]. While a detailed description is lacking regarding whether the biomarker measures were also extracted at the time of enrolment or not, it seems that there was the potential for the hs-CRP and D-dimer values to have been measured after the outcome. This highlights that even in a cross-sectional study, the recall or reference periods of measurements can introduce a temporal structure that should be considered on a variable-by-variable basis[7,29].

While cross-sectional studies, such as the novel work by Patil et al.[8], have an important role in epidemiology, an inability to establish a temporal ordering in the data does not mean that such considerations are unimportant[4,7]. Even if longitudinal data are not available and there is no attempt to estimate causal effects, models typically generate more informative estimates if there is at least a hypothesized temporality of the relationships between variables[3,4,7]. While the authors clearly articulate the limitations of their paper and conclude by saying that 'prospective studies are required to understand the relationship between food insecurity, hs-CRP and D-dimer better', I believe that they may have got more out of their efforts if Sir Austin Bradford Hill's question[1] was given more thought: which is the cart and which is the horse?

Acknowledgements

Taylor McLinden

British Columbia Centre for Excellence in HIV/AIDS
608–1081 Burrard Street
Vancouver, BC, Canada, V6Z 1Y6
Email: tmclinden@cfenet.ubc.ca

References

1. Hill AB (1965) The environment and disease: association or causation? *Proc R Soc Med* **58**, 295–300.
2. Ioannidis JP (2016) Exposure-wide epidemiology: revisiting Bradford Hill. *Stat Med* **35**, 1749–1762.
3. Rothman KJ & Greenland S (2005) Causation and causal inference in epidemiology. *Am J Public Health* **95**, Suppl. 1, S144–S150.
4. Phillips CV & Goodman KJ (2006) Causal criteria and counterfactuals; nothing more (or less) than scientific common sense. *Emerg Themes Epidemiol* **3**, 5.
5. Rogawski ET, Gray CL & Poole C (2016) An argument for renewed focus on epidemiology for public health. *Ann Epidemiol* **26**, 729–733.
6. Ho PM, Peterson PN & Masoudi FA (2008) Evaluating the evidence: is there a rigid hierarchy? *Circulation* **118**, 1675–1684.
7. Shahar E & Shahar DJ (2013) Causal diagrams and the cross-sectional study. *Clin Epidemiol* **5**, 57–65.

8. Patil S, Kadam D, Mehtani N *et al.* (2019) Elevated highly sensitive C-reactive protein and D-dimer levels are associated with food insecurity among people living with HIV in Pune, India. *Public Health Nutr.* Published online: 4 March 2019. doi: 10.1017/S136898001900020X.

9. Frega R, Duffy F, Rawat R *et al.* (2010) Food insecurity in the context of HIV/AIDS: a framework for a new era of programming. *Food Nutr Bull* **31**, issue 4, S292–S312.

10. McLinden T, Moodie EEM, Hamelin AM *et al.* (2017) Injection drug use, unemployment, and severe food insecurity among HIV–HCV co-infected individuals: a mediation analysis. *AIDS Behav* **21**, 3496–3505.

11. McLinden T, Moodie EEM, Harper S *et al.* (2018) Injection drug use, food insecurity, and HIV-HCV co-infection: a longitudinal cohort analysis. *AIDS Care* **30**, 1322–1328.

12. Anema A, Mehra D, Weiser SD *et al.* (2015) Drivers and consequences of food insecurity among illicit drug users. In *Health of HIV Infected People: Food, Nutrition and Lifestyle with Antiretroviral Drugs*, pp. 359–385. New York: Elsevier Publishing Inc.

13. McLinden T (2018) Injection drug use and food insecurity among HIV-hepatitis C virus co-infected individuals: associations, mechanisms, and interventions. PhD Thesis, McGill University.

14. Moodie EEM & Stephens DA (2010) Using directed acyclic graphs to detect limitations of traditional regression in longitudinal studies. *Int J Public Health* **55**, 701–703.

15. McLinden T, Moodie EEM, Hamelin AM *et al.* (2018) Methadone treatment, severe food insecurity, and HIV–HCV co-infection: a propensity score matching analysis. *Drug Alcohol Depend* **185**, 374–380.

16. Hanson KL & Connor LM (2014) Food insecurity and dietary quality in US adults and children: a systematic review. *Am J Clin Nutr* **100**, 684–692.

17. Kirkpatrick SI & Tarasuk V (2008) Food insecurity is associated with nutrient inadequacies among Canadian adults and adolescents. *J Nutr* **138**, 604–612.

18. Lee JS & Frongillo EA (2001) Nutritional and health consequences are associated with food insecurity among US elderly persons. *J Nutr* **131**, 1503–1509.

19. Mello JA, Gans KM, Risica PM *et al.* (2010) How is food insecurity associated with dietary behaviors? An analysis with low-income, ethnically diverse participants in a nutrition intervention study. *J Am Diet Assoc* **110**, 1906–1911.

20. Green MJ & Popham F (2019) Interpreting mutual adjustment for multiple indicators of socioeconomic position without committing mutual adjustment fallacies. *BMC Public Health* **19**, 10.

21. Westreich D & Greenland S (2013) The Table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *Am J Epidemiol* **177**, 292–298.

22. Bandoli G, Palmsten K, Chambers CD *et al.* (2018) Revisiting the Table 2 fallacy: a motivating example examining preeclampsia and preterm birth. *Paediatr Perinat Epidemiol* **32**, 390–397.

23. Greenland S & Pearce N (2015) Statistical foundations for model-based adjustments. *Annu Rev Public Health* **36**, 89–108.

24. Althoff KN, Smit M, Reiss P *et al.* (2016) HIV and ageing: improving quantity and quality of life. *Curr Opin HIV AIDS* **11**, 527–536.

25. Tesfaye M, Kaestel P, Olsen MF *et al.* (2016) Food insecurity, mental health and quality of life among people living with HIV commencing antiretroviral treatment in Ethiopia: a cross-sectional study. *Health Qual Life Outcomes* **14**, 37.

26. Sharkey JR, Johnson CM Dean WR (2011) Relationship of household food insecurity to health-related quality of life in a large sample of rural and urban women. *Women Health* **51**, 442–460.

27. Palermo T, Rawat R, Weiser SD *et al.* (2013) Food access and diet quality are associated with quality of life outcomes among HIV-infected individuals in Uganda. *PLoS One* **8**, e62353.

28. Casey PH, Szeto KL, Robbins JM *et al.* (2005) Child health-related quality of life and household food security. *Arch Pediatr Adolesc Med* **159**, 51–56.

29. Fedak KM, Bernal A, Capshaw ZA *et al.* (2015) Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol* **12**, 14.