Assessing bias in a prospective study of diabetes that implemented substitution sampling as a recruitment strategy

Michael C. David[a,*], Robert S. Ware[a], Rosa Alati[a], Jo Dower[b], Maria Donald[c]

[a]School of Population Health, The University of Queensland, Herston Road, Herston 4006, Queensland, Australia
[b]Health Quality and Complaints Commission, Level 17, 53 Albert Street, Brisbane 4000, Queensland, Australia
[c]Cancer and Population Studies Unit, The Queensland Institute of Medical Research, Herston, Road, Herston 4006, Queensland, Australia

Abstract

Objective: Strategies such as reminders are frequently used to maximize baseline recruitment and for this reason are collectively termed "usual practice." The objective of this study was to investigate substitution sampling as an alternative recruitment strategy.

Study Design and Settings: Data are from the Living with Diabetes Study, which is a prospective cohort study providing a comprehensive examination of health care utilization. Baseline information was collected for 3,197 of 11,470 eligible individuals between November 2008 and October 2009. Follow-up occurred 12 months after recruitment, with outcome of interest being emergency department attendance. Biases resulting from the two recruitment programs were investigated through the comparison of adjusted logistic regression coefficients and absolute relative biases (ARBs).

Results: Corresponding estimates resulting from both programs were similar except for age (75+ years). This effect was significant (β: -0.59; 95% confidence interval [CI]: -1.04, -0.13) under substitution sampling, but not under "usual practice" (β: -0.36; 95% CI: -0.78, 0.07). Analysis using the ARB metric reinforced similarity, with the Wilcoxon signed-rank test failing to detect significant difference between programs (median difference: -1.01; 95% CI: -5.88, 2.02).

Conclusion: Substitution sampling deserves consideration as a recruitment option alongside "usual practice," as concerns about additional bias may be unwarranted. © 2014 Elsevier Inc. All rights reserved.

Keywords: Bias; Prospective cohort study; Recruitment; Reminders; Responders; Substitution sampling

1. Introduction

A major threat to a study's external validity or generalizability is the recruitment of a nonrepresentative sample at baseline [1,2]. Endeavors to lessen this potential threat remain a methodological issue of ongoing concern. A multitude of strategies aimed to enhance baseline recruitment have been developed [3,4]. These strategies are rarely used in isolation. Rather, most studies rely on a combination of strategies [5−7]. The combined use of reminders and incentives can be considered the norm or usual baseline recruitment practice (forthwith called "usual practice"). Typically, reminders include cards, letters, replacement questionnaires, or telephone calls, with incentives being monetary or nonmonetary. However, the literature in support of this model, and thus the usage of reminders and incentives for nonresponder conversions is mixed [8−11].

Decisions regarding program make-up and implementation also need to consider more than statistical criteria. Strategies relating to "usual practice" can be costly and subject to diminishing returns [12,13]; consequently, programs incorporating new strategies that focus on increasing the recruitment of early responders have intuitive appeal. One such alternative strategy is substitution sampling, also known as booster, refreshment, and replacement sampling [14−16]. Substitution sampling involves the sourcing of a secondary group of potential responders from the original sampling frame, thus increasing the number of early responders by the use of two sampling stages. Usage of substitution sampling has been minimal to date, with support in the literature equivocal [14,17−19]. Underlying the lack of universal approval for substitution sampling is the concern that it has the potential to produce increased levels of bias [19] owing to its reliance on participants from only one side of the continuum of response spectrum [10,20].

The aim of the study was to investigate the levels of absolute relative bias (ARB) associated with substitution

---

* Corresponding author. Tel.: +61-07-3839-1621; fax: +61-07-3839-1631.

E-mail address: michael.david@uqconnect.edu.au (M.C. David).

What is new?

Key findings
- Substitution sampling need not produce more biased samples than those obtained from commonly used recruitment strategies such as reminders and incentives.

What this adds to what was known?
- Statistical inferences emanating from a sample of early responders (recruited through substitution sampling) were found to be comparable with corresponding inferences produced by a sample of early and late responders (recruited through "usual practice").

What is the implications and what should change now?
- When choosing recruitment strategies to enhance baseline response, consideration in addition to usual practice strategies, should be given to the use of substitution sampling.

---

sampling, as compared to "usual practice." Specifically, we hypothesized that a baseline recruitment program reliant on substitution sampling could produce estimates with comparable bias to those produced by a program of "usual practice." Data from an ongoing study focusing on people diagnosed with diabetes were used for investigative purposes.

2. Methods

2.1. The Living with Diabetes Study cohort

The Living with Diabetes Study (LWDS) is a prospective cohort study based in Queensland, Australia, involving adults diagnosed with diabetes, and has been described elsewhere [21,22]. Briefly, 14,439 adults were invited by postal questionnaire to participate in the study in November 2008. For the purposes of this specific study, an additional eligibility criterion pertaining to participant residence reduced this group to 9,970 (forthwith known as group 1 invitees). A secondary recruitment mail-out of 1,500 (forthwith known as group 2 invitees) was conducted in October 2009. The study was approved by the University of Queensland's Behavioural and Social Sciences Ethical Review Committee.

2.2. Data collection

To maximize participation from both groups, a number of strategies such as reminder cards and replacement

questionnaires were used. However, mail-out schedules for the two groups were not identical, with replacement questionnaires for group 1 being mailed 34 days after the original questionnaire and 101 days for group 2. Given time in transit was estimated to be a minimum of 1 week, respondents from group 1 were classified as early responders if the original questionnaire had been returned within 41 days. For group 2, early responders had an additional 20 days, as group 2 returns were not recorded as frequently as group 1 (ie, near-continuously for group 1, whereas block-based for group 2). Late responders for both groups were responders other than early responders.

Self-reported data on satisfaction with care, quality of life, health care utilization, and disease progression were obtained from the LWDS questionnaire, which consisted of more than 100 items. Variables of interest in this particular study were taken from a previous LWDS hospitalization study by Begum et al. [23]. The outcome variable of interest being emergency department attendance (EDA) in the past 12 months, and baseline variables being sex, age, marital status, level of education, income, disease duration, current depression, disease severity, and Patient Activation Measure (PAM).

2.3. Statistical analysis

Three response scenarios were defined, namely scenario 1 being the recruitment of all responders (early and late responders) from group 1 only, scenario 2 being the recruitment of early responders from groups 1 and 2 combined, and scenario 3 being the recruitment of all responders from groups 1 and 2 combined. Therefore, scenario 1 represented "usual practice," and scenario 2 represented recruitment by substitution sampling, whereas scenario 3 served as a proxy for the LWDS sampling frame. For each of the three scenarios and replicating the analysis by Begum et al. [23], univariable analyses were used to identify variables associated with EDA at a significance level of 0.10. Variables significant at this level and for which data were complete were then retained for multivariable logistic modeling. Owing to likely confounding, sex and age were automatically included.

To assess the extent to which substitution sampling (ie, scenario 2) might have introduced additional bias compared with "usual practice" (ie, scenario 1), adjusted regression coefficients (β) from each of these scenarios were compared with those from scenario 3. Further assessment was undertaken by comparing the ARBs of scenario 1 and 2 (both relative to scenario 3). This undertaking was both descriptive and analytical, with the latter using a Wilcoxon signed-rank test for small samples ($N < 15$) [24]. Scenario 3 was chosen as the proxy LWDS sampling frame for two reasons. First, individual-level data were limited to responders, meaning that responder vs. nonresponder comparisons could not be undertaken at the population level. Second and most important, the purpose of this study was