Participants Who Left a Multiple-Wave Cohort Study Had Similar Baseline Characteristics to Participants Who Returned

ROBERT S. WARE, PHD, GAIL M. WILLIAMS, PHD, AND ROSEMARY L. AIRD, BSOCSC(HONS)

PURPOSE: Research on determinants of an individual's pattern of response, considered as a profile across time, for cohort studies with multiple waves is limited. In this prospective population-based pregnancy cohort, we investigated baseline characteristics of participants after partitioning them according to their
history of response to different interview waves.
METHODS: Data are from the Mater-University of Queensland Study of Pregnancy 1981 to 1983 cohort, Brisbane, Australia. Complete baseline information was collected for 7223 of 7535 eligible individuals (95.9%). Follow-up occurred at 6 months, 5 years, and 14 years. Response rates were 93.0%, 72.5%, and
71.8%. Participants were allowed to leave and reenter the study. Participants were categorized as always, intermittent, or never responders. Intermittent responders were categorized further as leavers (responded at least once before leaving the study) or returners (left the study before reentering).
RESULTS: Participants who always responded were older, more educated, married, Caucasian, and nonsmokers and had higher incomes. Intermittent responders shared similar baseline characteristics. Relative risk for being an intermittent responder was located between risks for always or never responding.
CONCLUSIONS: Participants who left and reentered the study had baseline characteristics similar to participants who responded at least once and then left the study.
KEY WORDS: Bias, Cohort Study, Data Collection, Epidemiologic Methods, Health Surveys, Nonresponse, Questionnaires.

Introduction

Cohort studies are an important tool to examine causal relationships and specificity of exposure timing. Information on completeness of follow-up is crucial in assessing the generalizability and validity of a study [1,2]. Attrition, or loss of study participants, inevitably will lead to lower study power and less precise, possibly biased, estimates [3,4]. Comparisons between respondents and nonrespondents are reported routinely in published studies[5, 6, 7, 8,9]. However, most multiple-wave prospective studies report only nonresponse at a single wave [10, 11, 12]. Little information is available on factors associated with an individual participant's history of response across multiple study waves, especially for studies in which participants are permitted to leave and reenter.

This report documents results of a 14-year follow-up survey of an Australian pregnancy cohort, the Mater-University of Queensland Study of Pregnancy, with almost complete information at entry and multiple follow-ups. Our study is designed to test the hypothesis that participants who leave and reenter potentially are informative about outcomes of participants who respond at least once before leaving the study and are distinct from participants who always respond and participants who never respond.

Methods

The Mater-University of Queensland Study of Pregnancy

The Mater-University of Queensland Study of Pregnancy, a cohort study focusing on assessment of mother and child health and social outcomes, was described previously[13]. Briefly, a woman was eligible to enter the study if she attended her first hospital obstetric visit at, delivered at, and was discharged with a live singleton birth from the Mater Misericordiae Hospital, Brisbane, Australia. Enrollment took place between January 1981 and December 1983 and was restricted to public patients, that is, women who attended a free hospital clinic for antenatal care and were delivered under the supervision of full-time medical staff. (During 1981 to 1983, a total of 46% of all deliveries at Mater Misericordiae Hospital were public.) Of 7535 women who met enrollment criteria, 7223 (95.9%) agreed to be interviewed 3 to 5 days after delivery. Information was collected on demographic and socioeconomic characteristics, health behavior, and health status. These individuals are defined as participants for the purpose of the study and are targeted in future follow-ups.

Follow-Up Surveys

Follow-up waves were conducted 6 months, 5 years, and 14 years after birth. Data were collected on maternal demographics, lifestyle, mental health, child rearing and health, pregnancy outcomes, child health, and child behavior. There was no interim contact with participants.

At each follow-up, participants were recontacted through telephone and/or address details they had provided at the previous wave (including contact details of up to four relatives or friends). During both the 5- and 14-year follow-ups, participants were invited to attend an interview at the study hospital. If the participant could not attend the interview, the questionnaire was mailed to them. For those who agreed to an interview, but were unable to travel to the study hospital, interviews were conducted in their homes. Any participant who actively withdrew from the study was not recontacted. Ethics approval was gained from relevant committees at The University of Queensland and Mater Misericordiae Hospital.

Participants were partitioned into four groups according to history of response at the 6-month, 5-year, and 14-year follow-ups. Response is defined as a mother or caregiver completing a questionnaire. *Always responders* are defined as participants who responded at all three follow-up waves. *Never responders* are participants who did not respond at any of the three follow-up waves. Intermittent responders were split into two groups; *returners* are participants who missed at least one wave before returning to the study to take part in the 14-year follow-up, whereas *leavers* are participants who responded at one or both of the 6-month and 5-year follow-ups, but did not respond at the 14-year follow-up.

Statistical Analyses

To find variables associated most strongly with attrition, a number of variables measuring key demographic (age, education, employment, marital status, income, and ethnicity), health behavior (smoking status, alcohol consumption, and level of physical activity), and mental health (depression and anxiety) criteria, as recorded at baseline, were investigated. A variable was considered for the multivariate model if it was significant at the $p = 0.1$ level in univariate logistic regression with response (yes/no) as the outcome. No adjustment was made for multiple comparisons. Using multivariable logistic regression with response (yes/no) as the

outcome, we identified the most influential variables and investigated potential interactions between variables. Polytomous logistic regression was used to simultaneously compare returners, leavers, and never responders with always responders. Relative risks (RRs) for response with 95% confidence intervals (CIs) were calculated, using always responders as the reference category.

Results

There were 6720 participants (93.0%) who responded at the 6-month follow-up; 5234 (72.5%), at the 5-year follow-up; and 5185 (71.8%), at the 14-year follow-up. Age, education, income, marital status, ethnicity, smoking, level of physical activity, and anxiety, all measured at baseline, were identified as the variables most strongly associated with response. No interaction terms were statistically significant. Response rates at each of the three follow-up waves were greater for participants who were older, more educated, wealthier, married, Caucasian, nonsmokers, physically active, and not anxious. Participants with missing data items at baseline were more likely not to respond at each wave than participants with nonmissing items (data not shown).

Baseline Characteristics of Participants Partitioned by History of Response to Different Interview Waves

There were 4470 (61.9%) always responders, 1747 (24.2%) leavers, 715 (9.9%) returners, and 291 (4.0%) never responders (Table 1). The four response history groups were statistically significantly different at the $p = 0.001$ level for all characteristics. However, there were no significant differences between leavers and returners at $p = 0.05$ (unadjusted data, chi-squared test). After adjustment for all identified variables, the group with the RR for response most different from always responders (the reference group) was never responders. Generally, returners and leavers had similar RRs for response; these risks were located between the always and never responders. For example, the RR for smoking at baseline was 1.68 (95% CI, 1.29–2.20) for never responders relative to always responders, with returners and leavers in between (RR, 1.36; 95% CI, 1.15–1.62) and (RR, 1.39; 95% CI, 1.23–1.57), respectively.

Table 1. Relative risk for being in a subgroup by history of response to different interview waves relative to always responders, Mater-University of Queensland Study of Pregnancy, Brisbane, Australia 1981 to 1983

Characteristic	N	Crude Relative Risk (95% CI)			Adjusted[a] Relative Risk (95% CI)		
		Returner	Leaver	Never Responder	Returner	Leaver	Never Responder
Age (years)							
13–19	1181	1.00	1.00	1.00	1.00	1.00	1.00
20–29	4802	0.48 (0.40–0.59)	0.54 (0.46–0.62)	0.35 (0.27–0.47)	0.63 (0.51–0.79)	0.75 (0.64–0.89)	0.65 (0.48–0.89)
30–49	1240	0.42 (0.32–0.55)	0.49 (0.40–0.59)	0.23 (0.15–0.35)	0.58 (0.43–0.78)	0.70 (0.57–0.87)	0.46 (0.29–0.75)
Education (years completed)							
≤12	5914	1.00	1.00	1.00	1.00	1.00	1.00
≥13	1256	0.67 (0.53–0.84)	0.79 (0.68–0.91)	0.40 (0.27–0.61)	0.77 (0.61–0.98)	0.90 (0.77–1.05)	0.54 (0.35–0.82)
Household income[b]							
Low	2308	1.00	1.00	1.00	1.00	1.00	1.00
High	4441	0.60 (0.51–0.71)	0.54 (0.47–0.60)	0.31 (0.24–0.40)	0.77 (0.65–0.93)	0.69 (0.60–0.78)	0.46 (0.35–0.61)
Marital status							
Married, cohabitant	5386	1.00	1.00	1.00	1.00	1.00	1.00
Unmarried, cohabitant	844	2.25 (1.78–2.84)	2.49 (2.11–2.95)	4.69 (3.47–6.36)	1.68 (1.31–2.15)	1.97 (1.65–2.36)	3.00 (2.15–4.16)
Unmarried, not cohabitant	736	2.09 (1.63–2.67)	2.21 (1.85–2.64)	3.91 (2.82–5.43)	1.30 (0.98–1.72)	1.41 (1.15–1.73)	1.86 (1.27–2.71)
Other	194	3.02 (1.96–4.66)	3.10 (2.24–4.31)	4.88 (2.74–8.70)	2.42 (1.55–3.79)	2.37 (1.69–3.33)	3.04 (1.66–5.56)
Race mother							
Caucasian	6443	1.00	1.00	1.00	1.00	1.00	1.00
Aboriginal/Islander	289	3.18 (2.23–4.55)	3.10 (2.36–4.09)	4.94 (3.20–7.61)	2.70 (1.88–3.89)	2.52 (1.90–3.35)	3.51 (2.24–5.52)
Other	249	1.30 (0.82–2.04)	2.13 (1.61–2.81)	1.84 (1.03–3.32)	1.64 (1.03–2.63)	2.63 (1.96–3.51)	2.65 (1.41–4.98)
Smoking							
Nonsmoker	3623	1.00	1.00	1.00	1.00	1.00	1.00
Current smoker	3546	1.63 (1.39–1.91)	1.63 (1.45–1.82)	2.28 (1.78–2.93)	1.36 (1.15–1.62)	1.39 (1.23–1.57)	1.68 (1.29–2.20)
Physical activity[c]							
Physically active	6437	1.00	1.00	1.00	1.00	1.00	1.00
Not physically active	653	1.34 (1.03–1.75)	1.32 (1.09–1.59)	1.84 (1.29–2.61)	1.36 (1.03–1.78)	1.29 (1.06–1.58)	1.85 (1.28–2.68)
Maternal anxiety[d]							
Not anxious	q	1.00	1.00	1.00	1.00	1.00	1.00
Anxious	926	1.65 (1.33–2.06)	1.72 (1.47–2.02)	1.97 (1.44–2.69)	1.31 (1.05–1.65)	1.35 (1.14–1.59)	1.31 (0.95–1.81)

Characteristics measured at baseline. Numbers in each group: always responders, 4470; returners, 715; leavers, 1747; and never responders, 291.
[a] Estimates adjusted for all identified variables.
[b] Low household income, less than $10,400 Australian dollars per annum (1981 to 1983 value).
[c] Never: response to the question, "Before you were pregnant, how often did you undertake physical exercise or active sports?" was "never."
[d] Anxious: scored 4 or higher on the seven-item anxiety scale from the Delusion Symptom States Inventory[14].

Discussion

This report provides response rates for a 14-year pregnancy-based cohort that has almost complete information at entry and multiple follow-ups, with participants permitted to leave and reenter the study. Participants who missed one or more follow-up waves were willing to return to the study and participate at future follow-ups. By characterizing participants by their history of response to different interview waves, a clearer pattern of attrition has emerged. Always responders were found to be significantly statistically different from never responders. Leavers and returners share similar baseline characteristics.

Results of this study have practical application for researchers in two ways. First, this information may be used to keep individuals with a high risk for dropping out in the study, for example, by taking additional measures to improve participation rates in these groups. Second, if researchers choose to adjust for missing data by using imputation or modeling, they need to make assumptions about the characteristics of subjects who do not respond. By considering the difference in characteristics by response history, there is the potential to conduct more sophisticated analyses.

One limitation of our study is that we are not able to distinguish between reasons for attrition at each follow-up wave. When different types of attrition are properly distinguished, a clearer pattern of attrition should emerge[15,16]. Four major classes of attrition have been identified in the literature: an eligible participant declines to enter the study at its initialization, a participant declines to complete a questionnaire at a subsequent follow-up, there is nonresponse due to incapacity or death, or researchers fail to relocate a participant for a follow-up wave. Previously, large epidemiologic studies found different predictors for each of these reasons for attrition[17,18]. The design of the Mater-University of Queensland Study of Pregnancy did not allow us to easily distinguish whether nonresponse at a follow-up wave was caused by participant refusal, researchers not being able to relocate a participant, or because of participant incapacity. We expect the number of participants who did not respond because of death or incapacity to be relatively small because we are dealing with a 14-year follow-up of new mothers. Nevertheless, response rates may be understated because ineligible individuals are included[19]. Ideally, sample lists would have been corrected before data collection; however, in the current study, this was not possible, and to make adjustments post hoc, we would need to ascertain the eligibility status of survey nonresponders.

In summary, a select group of individuals returned for follow-up. When partitioned by history of response to different interview waves, there were significant differences between individuals who always, intermittently, and never responded. The two types of intermittent responders shared similar baseline characteristics. These findings should be considered when addressing potential biases introduced by nonresponding participants and when determining which participants should be specifically targeted to curb participant losses in follow-up waves in long-term prospective epidemiologic studies.

References

1 S. Greenland, Response and follow-up bias in cohort studies, Am J Epidemiol 106 (1977), pp. 184–187.

2 J.H.D. Deeg, Attrition in longitudinal population studies: Does it affect the generalizability of the findings? An introduction to the series, J Clin Epidemiol 55 (2002), pp. 213–215.

3 O.S. Miettinen, Theoretical Epidemiology: Principles of Occurrence Research in Medicine, Wiley, New York (1985).

4 K. Rothman, Epidemiology: An Introduction, Oxford University, New York (2002).

5 M.H. Criqui, E. Barrett-Connor and M. Austin, Differences between respondents and nonrespondents in a population-based cardiovascular disease study, Am J Epidemiol 108 (1978), pp. 367–372.

6 R.C. Klesges, J.E. Williamson, G.W. Somes, G.W. Talcott, H.A. Lando and C.K. Haddock, A population comparison of participants and nonparticipants in a health survey, Am J Public Health 89 (1999), pp. 1228–1231.

7 L.K. Heilbrun, A. Nomura and G.N. Stemmermann, The effects of nonresponse in a prospective study of cancer, Am J Epidemiol 116 (1982), pp. 353–363.

8 D.M. Purdie, M.P. Dunne, F.M. Boyle, M.D. Cook and J.M. Najman, Health and demographic characteristics of respondents in an Australian national sexuality survey: Comparison with population norms, J Epidemiol Community Health 56 (2002), pp. 748–753.

9 R. Jackson, L.E. Cambless, K. Yang, T. Byrne, R. Watson and A. Folsom et al., Differences between respondents and nonrespondents in a multicenter-community-based study vary by gender and ethnicity, J Clin Epidemiol 49 (1996), pp. 1441–1446.

10 A. Britton, A. Singh-Manoux and M. Marmot, Alcohol consumption and cognitive function in the Whitehall II study, Am J Epidemiol 160 (2004), pp. 240–247.

11 P.M. Eng, E.B. Rimm, G. Fitzmaurice and I. Kawachi, Social ties and change in social ties in relation to subsequent total and cause-specific mortality and coronary heart disease incidence in men, Am J Epidemiol 155 (2002), pp. 700–709.

12 L.F. Berkman, M. Melchior, J.-F. Chastang, I. Niedhammer, A. Leclerc and M. Goldberg, Social integration and mortality: A prospective study of French employees of Electricity of France-Gas of France, Am J Epidemiol 159 (2004), pp. 167–174.

13 J.D. Keeping, J.M. Najman, J. Morrison, J.S. Western, M.J. Andersen and G.M. Williams, A prospective longitudinal study of social, psychological and obstetric factors in pregnancy: Response rates and demographic characteristics of the 8556 respondents, Br J Obstet Gynecol 96 (1989), pp. 289–297.

14 A. Bedford, G.A. Foulds and B.F. Sheffield, A new personal disturbance scale (DSSI/sAD), Br J Soc Clin Psychol 15 (1976), pp. 387–394.

Ann Epidemiol 2006;16:820–823.

15 C.E.M. Van Beijsterveldt, M.P.J. van Boxtel, H. Bosma, P.J. Houx, F. Buntinx and J. Jolles, Predictors of attrition in a longitudinal cognitive aging study: The Maastrict Aging Study (MAAS), J Clin Epidemiol 55 (2002), pp. 216–223.

16 G.I.J.M. Kempen and E. van Sonderen, Psychological attributes and changes in disability among low-functioning older persons: Does attrition affect the outcomes?, J Clin Epidemiol 55 (2002), pp. 224–229.

17 M.E. Framer, B.Z. Locke, I.Y. Liu and E.K. Moscicki, Depressive symptoms and attrition: The NHANES I epidemiologic follow-up study, Int J Methods Psychiatr Res 4 (1994), pp. 19–27.

18 W.W. Eaton, J.C. Anthony, S. Tepper and A. Dryman, Psychopathology and attrition in the Epidemiologic Catchment Area Surveys, Am J Epidemiol 135 (1992), pp. 1051–1059.

19 B.J. Harvey, A.L. Wilkins, G.A. Hawker, E.M. Badley, P.C. Coyte and R.H. Glazier et al., Using publicly available directories to trace survey nonresponders and calculate adjusted response rates, Am J Epidemiol 158 (2003), pp. 1007–1011.