

Feasibility and Acceptability of Alternate Methods of Postnatal Data Collection

Lacey A. McCormack · Christa Friedrich ·
Nancy Fahrenwald · Bonny Specker

Published online: 22 June 2013
© Springer Science+Business Media New York 2013

Abstract This study was done in preparation for the launch of the National Children’s Study (NCS) main study. The goal of this study was to examine the feasibility (completion rates and completeness of data), acceptability, staff time and cost-effectiveness of three methods of data collection for the postnatal 3- and 9-month questionnaires completed as part of NCS protocol. Eligible NCS participants who were scheduled to complete a postnatal questionnaire at three and nine months were randomly assigned to receive either: (a) telephone data collection (b) web-based data collection, or (c) self-administered (mailed) questionnaires. Event completion rates and satisfaction across the three data collection methods were compared and the influence of socio-demographic factors on completion rates and satisfaction rates was examined. Cost data were compared to data for completion and satisfaction for each of the delivery methods. Completion rates and satisfaction did not differ significantly by method, but completeness of data did, with odds of data completeness higher among web than phone ($p < 0.001$) or mail ($p < 0.001$). Costs were highest for the phone, followed by mail and web methods ($p < 0.001$). No significant differences in participant time (i.e. burden) across the three data collection methods were seen. Mail and phone data collection were the least complete of the three methods and were the most expensive. Mailed data collection was neither complete nor exceptionally economical. Web-based data collection was the least costly and provided the most

complete data. Participants without web access could complete the questionnaire over the phone.

Keywords Data collection/methods · Feasibility · Women · Questionnaires

Introduction

Participation in epidemiologic studies in general has been declining over past decades [1]. A decline in participation has been documented in the administration of telephone surveys as well [2]. Declining participation, whether through poor study recruitment or decreased response rates to questionnaires, can impact data quality. This issue of nonresponse may especially be a problem for post-partum women who are adjusting to life with a new baby, so importance should be placed upon determining the ideal method of data collection for this population. While a few studies have assessed response rates in web versus mailed/paper surveys [3–8] and web versus telephone surveys [9–11] very little research exists that compares the feasibility, acceptability and cost-effectiveness of telephone, web-based and mailed/paper survey data collection. The few studies that do exist focused on health surveillance and patient satisfaction and highlight the need for additional research in this area. Findings from one prior study indicate that demographics vary considerably for respondents to web and mailed surveys compared to telephone respondents [12]. Mailed surveys seem to be a viable alternative to telephone surveys, while web-based surveys can be problematic when it comes to gathering information from study participants due to nonresponse bias [13]. E-mail has been cited as the most cost-efficient method of survey administration [14]. We are unaware of any research that

L. A. McCormack (✉) · C. Friedrich · N. Fahrenwald ·
B. Specker
EA Martin Program in Human Nutrition, South Dakota State
University NCS Study Center, SWC, Box 506, Brookings,
SD 57007, USA
e-mail: lacey.mccormack@sdstate.edu

exists identifying the ideal method of questionnaire data collection in post-partum women aged 18–49. This study was done in preparation for the launch of the main study for the National Children’s Study (NCS), a nationwide longitudinal cohort study that will examine environmental influences on child health and development from preconception or birth until 21 years of age. Women between the ages of 18–49 were recruited for the NCS through enumeration of eligible households and pregnancy screening. A consented woman participated in data collection throughout pregnancy and at selected time points after the child was born. Additional details about the study can be found elsewhere [15].

Understanding the ideal way for collecting survey data in a lengthy longitudinal study like the NCS is important, therefore the goal of this study was to examine the feasibility (completion rates and completeness of data), acceptability, staff time and cost-effectiveness of three methods of data collection for the postnatal 3- and 9-month questionnaires (PN3 and PN9) completed as part of NCS protocol. The PN3 and PN9 questionnaires collect information on a variety of topics ranging from sleeping habits, crying patterns, developmental milestones, childcare and doctor visits in baby to doctor visits, hospitalizations and spousal relationships in mom, similar to those questions utilized in pregnancy risk assessment monitoring system (PRAMS) research [16].

Methods

To determine study eligibility, a brief screening questionnaire was created to assess participants’ ability to complete each of the three data collection methods (web, mail and telephone). Those women who indicated they were not able to participate in data collection via one or more of the methods were not eligible. A web-based questionnaire was created from the original, paper-based phone questionnaire, as was a version suitable for mailing. Additionally, a short satisfaction survey was developed asking if the assigned data collection method was simple, convenient and would they choose it over the other two methods. A final question asked the participant which method of data collection out of the three they would prefer. The intention was to get someone’s initial thoughts and then see if preferences changed once the data collection method had actually been used. For example, did someone originally say they would prefer mail but later decide they prefer online? Or did someone originally think mail was easier to complete than web or phone but later change their mind? Finally, a cost review summary sheet was designed to track the personnel time and resources required for the three methods of data collection. Office of Management and Budget approval

(required for NCS formative research) and South Dakota State University Institutional Review Board approval were obtained prior to data collection.

Current NCS participants who were scheduled to complete a postnatal questionnaire at three or nine months (PN3 or PN9 events) and were eligible for the project, were randomly assigned to receive either: (a) telephone data collection according to current NCS protocol (b) web-based data collection, or (c) self-administered (mailed) questionnaires. Data were collected on events that opened between August and December of 2011. Due to the timing of the study, a woman may have completed just a PN3, just a PN9 or both as part of this project. Events needed to be completed within a certain timeframe around certain dates which were calculated based on a child’s birth date. The window for the PN3 event completion was 2–4 months post-partum and the window for the PN9 event was 8–10 months post-partum. Protocol restricted completion of the questionnaire to mothers. Both English and Spanish language versions of data collection were available. All data collection methods had the potential to be completed over multiple days and each allowed the participant to skip any question she did not want to answer. Phone interviews were completed according to current protocol, which involved a trained data collector calling the participant to administer the questionnaire over the phone during the event window. Various days of the week and times of day were tried if the participant was difficult to reach. The maximum number of call attempts per participant was up to the discretion of the data collector. For the mailed questionnaire group, when the event window opened, paper-based interviews were sent to the participants along with an explanation of the questionnaire. The participant received a reminder call 1 week before the event closing date if the site office had not yet received the paper questionnaire. To capture the amount of time it took the participant to complete the questionnaire, boxes were included on the questionnaire itself for the participant to indicate start and end time. Participants in the electronic data collection group were e-mailed a link to the online questionnaire along with an explanation of the link when their event window opened. If no e-mail address was available, or if the email was not delivered, the participants were mailed an event summary that included a link to the web-based questionnaire. The participant received a reminder call 1 week before the closing date if the site office had not received the web-based data collection responses. Upon project completion, all forms from all data entry methods were entered into a separate electronic database designed for the project to allow for dataset creation and comparison. For each data collection method, after the questionnaire was completed the participant was mailed a small monetary incentive. The incentive was the same for each

data collection method, and the participant was made aware of the incentive at a prior study visit.

Data Analysis

Data collection for this project occurred at the South Dakota State University Study Center and included participants from Brookings County, SD and Yellow Medicine, Pipestone and Lincoln counties in Minnesota. Data obtained from the three PN3 and PN9 data collection methods were reviewed by two study personnel to determine completeness of the data. Questions that were missing were coded with an 'X' and the event was assigned an 'Incomplete' status. The three satisfaction questions (scored 1–4) were averaged and those responses with a mean score of 3 or higher were considered satisfactory. Demographic data were obtained at previous study events.

Socio-demographic factors of participants who were eligible for the project were compared with participants who were not, using contingency table analysis and logistic regression. Event completion rates and satisfaction across the three data collection methods were compared using contingency table analysis (satisfactory or unsatisfactory) and the influence of socio-demographic factors such as age, race/ethnicity, education and income on completion rates and satisfaction rates was examined using logistic regression. Cost of staff time and resources (postage, printing, web-site maintenance) required for each data collection method was determined. Of note, phone costs do not include charges for individual phone lines or long distance charges. Additionally, staff time does not include that for data entry, as this time was part of the project (for ease of data comparison). Cost data were compared to data for completion (feasibility) and satisfaction (acceptability) for each of the delivery methods. Additionally, completion rates were calculated from the number of incomplete events open and the number of events completed. In regards to reporting average time for event completion, events were excluded that spanned multiple days ($n = 8$) as they indicated the overall length of time for questionnaire completion from start to finish but not the actual time spent completing the questionnaire.

Results

The NCS Program Office has guidelines surrounding participant disclosure, so numbers <10 have not been reported. Overall, 238 eligibility questionnaires were returned, with 41 participants (17 %) indicating that they could not complete data collection via one or more of the methods. Demographic differences between the eligible and ineligible groups are shown in Table 1. Eligibility differed

significantly by marital status, income level, education level and ethnicity, with higher odds of eligibility among women who were married, had a college education, who were not Hispanic or Latino and whose household income level was \$75,000 or more versus \$30,000–\$74,999. Ultimately, 197 individuals (83 %) were eligible for the project, with a total of 107 upcoming PN3 or PN9 events for 97 unique participants during the project timeframe.

Data were successfully collected via all three methods for 93 of the 107 events (87 %), with 11 of the incomplete being PN3 data collection points (number of missing items per questionnaire ranged from 0–13). Of the incomplete PN3 events, most were web and mail, followed by phone. Of the few remaining incomplete PN9 events, most were mail followed by web. Randomization was successful as no demographic differences among the three data collection groups were found. Additionally, no significant differences in demographic information existed between participants who completed an event and participants who did not. Completion rates and satisfaction did not differ significantly by method, but completeness of data did (Table 2), with odds of data completeness higher among web than phone ($p < 0.001$) or mail ($p < 0.001$) data collection. Among those participants randomized to mailed questionnaires, no socio-demographic factors that were independent predictors of data completeness were found. Completeness for web-based data collection was not assessed due to lack of incomplete data, and phone data were obtained by data collectors, so participant demographics in regard to data completeness were irrelevant.

Average time to complete an event and cost by the different data collection methods is shown in Table 3. Costs were highest for the phone method, followed by mail and web methods ($p < 0.001$). No significant differences in participant time (i.e. burden) among the three data collection methods existed. Costs for web and mail methods were essentially fixed at \$5.44 and \$12.41 per completed questionnaire, respectively, as they required minimal staff time and were fixed materials and postage costs. Costs for the phone method ranged from \$10.91 to \$24.35 per completed questionnaire, depending on the length of the questionnaire administration (average = \$14.63 per completed questionnaire).

Discussion

The purpose of this study was to examine the feasibility (completion rates and completeness of data), acceptability, staff time and cost-effectiveness of three methods of data collection for questionnaires administered to post-partum women as part of the NCS. A few studies have examined these concepts in the context of broad survey

Table 1 Characteristics of total potential study population

	Eligible ^c	Ineligible ^c	OR (CI), <i>p</i> value
Eligible/ineligible	197	41	
Age (mean ± SD) ^a	30.1 ± 4.7	28.5 ± 4.8	NS
Marital status ^b			
Married	175 (89 %)	31 (76 %)	3.0 (1.2–7.2), 0.02
Not married	17 (8 %)	–	1.0
Household income level ^b			
Less than \$30,000	27 (14 %)	–	1.1 (0.4–2.9), NS
\$30,000–74,999	97 (49 %)	27 (66 %)	1.0
\$75,000 or more	63 (32 %)	–	2.9 (1.2–8.2), 0.02
Education level ^b			
Some high school, high school diploma or some college	43 (22 %)	24 (59 %)	1.0
College degree or higher	112 (57 %)	12 (29 %)	5.2 (2.4–11.7), < 0.001
Race ^b			
White	163 (83 %)	31 (76 %)	2.1 (0.8–4.8), NS
Non-White	23 (12 %)	–	1.0
Ethnicity ^b			
Not Hispanic or Latino	183 (93 %)	36 (88 %)	6.8 (1.4–35.6), 0.02
Hispanic or Latino	–	–	1.0
Primary language ^b			
English	193 (98 %)	37 (90 %)	10.4 (1.0–227.7), NS
Non-English	–	–	1.0

– Number too small (<10) to be displayed per NCS Program Office guidelines

^a Comparisons for pair using Student’s *t*

^b Differences between groups were analyzed using nominal logistic regression and odds ratios with confidence intervals

^c Numbers may not add up to group totals (and percents may not add up to 100) due to missing or incomplete information

Table 2 Completion rates, data completeness and satisfaction among data collection methods

Total events	Web ^a n = 37 (%)	Mail ^a n = 36 (%)	Phone ^a n = 34 (%)	<i>p</i> value ^b
Completion rates (% of total)	31 (84)	29 (81)	33 (97)	0.06
Complete (no missing data)	28 (90) ^{bc}	14 (48) ^b	15 (45) ^c	<0.001 ^c
Satisfactory (method was satisfactory)	16 (52)	17 (59)	17 (52)	NS

^a Numbers may not add up to group totals due to missing or incomplete information

^b All data analyzed using Likelihood Ratio Chi Square

^c Differences among groups were analyzed using nominal logistic regression; values with similar superscripts are different at *p* < 0.05

administration; however the impact of data collection method within an epidemiological study with regular participant contact is unknown. Link and colleagues have examined effects of mail, web and telephone data collection methods on self-reported alcohol consumption and found that method impacted survey response [13]. Specifically, individuals who responded via web were more likely to report heavy drinking than individuals who responded via mail or telephone. Although how responses were impacted in our study is unknown, completeness of data varied by method, with web-based data collection being the most complete. This difference is likely due to the participant being guided through questions automatically, where both the telephone and mailed methods were completed on

paper leaving more room for human error due to complicated skip patterns. Having the web-based method be the most complete method could indicate a participant’s willingness to share more information this way.

Additionally, endoscopy patient satisfaction was examined using mail, e-mail and telephone methods and response rates were highest among phone, mail and e-mail, respectively [14]. In the present study, 87 % of events with open windows were completed, regardless of method, with phone also having the highest completion rate (97 %), followed by web (84 %) and mail (81 %). The highest number of missed events (n = 11, 79 %) occurred for the PN3 event while only 3 (21 %) PN9 events were missed. The window for completing the PN3 event was 2–4 months

Table 3 Average cost per completed questionnaire for each data collection method

	Web	Mail	Phone	<i>p</i> value ^b
Participant time ^a	38 min	41 min	26 min	NS
Staff time	10 min (preparation)	10 min (preparation)	26 min + 10 min (preparation)	–
Staff cost (28.6 cents/min)	\$2.86	\$2.86	\$10.30	–
Paper/Printing cost	–	\$1.75	\$1.75	–
Mailing cost ^c	\$2.58	\$7.80	\$2.58	–
Total cost	\$5.44 ^{b,c}	\$12.41 ^{a,b}	\$14.63 ^{b,c}	<0.001

^a Average time; excludes events that were completed over the course of more than one day ($n = 8$)

^b *p* values determined by one-way ANOVA; means with similar superscripts are different at $p < 0.05$ (Tukey HSD for multiple comparisons)

^c Includes mailing of incentive upon event completion

post-birth, which can be an exceptionally busy and stressful time for parents. Interestingly, 93 % of the PN3 events that were missed were self-administered methods, indicating that data collection methods like telephone, where the participant burden is the smallest and completion is initiated by a data collector, are less likely to be missed. Likewise, all missed PN9 events were self-administered. The number of missed PN9 events was lower than that of the PN3 events, possibly indicating participants having more time (perhaps because of more established routines or fewer distractions) for data collection events 8–10 months after birth than 2–4 months after.

When examining cost-effectiveness of a mailed, phone and internet survey in Australia (with phone and internet having a postcard precede it), Sinclair et al. [17], found the lowest cost for generic mailed surveys. In our study, web-based data collection was the cheapest, with minimal staff time and costs for mailing incentives. Mailed data collection was more expensive even though it also required minimal staff time, due to the large number of forms that needed to be printed and mailed to the participant with return postage paid. In our study, information about the web-based questionnaire was e-mailed to the participant versus mailed, therefore reducing costs, which could explain the discrepancy in findings when it comes to cost.

The few studies done in this area have pointed toward the need for a mixed approach to survey administration, leveraging resources with participant preferences. Although cheaper than the phone method, both web-based and mailed data collection methods were self-administered and comprised 93 % of the 14 incomplete events. Event completion via phone demanded the least participant burden in terms of interview completion time and participant resources—accessing the internet and typing out responses wasn't required, nor was filling out multiple pages of forms by hand. Using phone data collection, the participant simply needed to be available to answer questions. All three methods provided flexibility however, and were able to be completed over multiple days if needed.

Limitations

The findings of this study are partially limited by the established research relationships that existed with study participants prior to this study. Participants had previous face to face, telephone and mailed questionnaire contact with study personnel, perhaps contributing to a greater completion rate across all data collection methods. Furthermore, the NCS population being followed by the SDSU Study Center is not representative of all NCS participants; however characteristics may be similar to other rural study locations in terms of access for different methods of data collection. Therefore, similar results might not be seen across the entire study. Additionally, the results are restricted to only women who were eligible for the project, further limiting generalizability. Finally, no charges were associated with web-based data collection itself for this study, as access to an online survey tool already existed. Costs associated with the purchase of such a tool for other studies could alter the cost-effectiveness of the web-based method and would need to be further examined.

Conclusions

Although actual completion rates were highest among the phone group, mail and phone data collection were the least complete of the three methods in terms of questionnaires with missing data, and they were the most expensive. Given the lack of difference in satisfaction among the three methods, it could make sense to start data collection with the least costly and most complete method—which in this study was web-based data collection. If a participant is unwilling or unable to complete the questionnaire via the web, or the end of the visit window is being reached, an alternative would be to complete over the phone. This combination would combine the strengths of both methods while keeping costs down. Mailed data collection, while not exceptionally complete nor economical, would be a last

resort data collection method for those participants unreachable by phone.

Acknowledgments This analysis was conducted as part of the National Children’s Study, supported by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, and funded, through its appropriation, by the Office of the Director of the National Institutes of Health. Supported in part by NICHD Contract Numbers HHSN275200603416C and HHSN275201100004C. The views expressed in this article are the responsibility of the authors and do not necessarily represent the position of the National Children’s Study, the National Institutes of Health, or the U.S. Department of Health and Human Services.

References

- Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of Epidemiology*, *17*(9), 643–653.
- Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, *69*(1), 87–98.
- Kroth, P. J., McPherson, L., Leverence, R., et al. (2009). Combining web-based and mail surveys improves response rates: A PBRN study from PRIME Net. *Annals of Family Medicine*, *7*(3), 245–248.
- Lusk, C., Delclos, G. L., Burau, K., Drawhorn, D. D., & Aday, L. A. (2007). Mail versus internet surveys: Determinants of method of response preferences among health professionals. *Evaluation and the Health Professions*, *30*(2), 186–201.
- McCabe, S. E., Couper, M. P., Cranford, J. A., & Boyd, C. J. (2006). Comparison of Web and mail surveys for studying secondary consequences associated with substance use: Evidence for minimal mode effects. *Addictive Behaviors*, *31*(1), 162–168.
- McCabe, S. E., Diez, A., Boyd, C. J., Nelson, T. F., & Weitzman, E. R. (2006). Comparing web and mail responses in a mixed mode survey in college alcohol use research. *Addictive Behaviors*, *31*(9), 1619–1627.
- Ritter P, Lorig K, Laurent D, Matthews K. (2004) Internet versus mailed questionnaires: A randomized comparison. *Journal of Medical Internet Research*. Sep 15 *6*(3):e29.
- Yetter, G., & Capaccioli, K. (2010). Differences in responses to Web and paper surveys among school professionals. *Behavior Research Methods*, *42*(1), 266–272.
- Graham, A. L., & Papandonatos, G. D. (2008). Reliability of internet-versus telephone-administered questionnaires in a diverse sample of smokers. *Journal Of Medical Internet Research*, *10*(1), e8.
- Greene, J., Speizer, H., & Wiitala, W. (2008). Telephone and web: Mixed-mode challenge. *Health Services Research*, *43*(1 Pt 1), 230–248.
- Rankin, K. M., Rauscher, G. H., McCarthy, B., et al. (2008). Comparing the reliability of responses to telephone-administered versus self-administered Web-based surveys in a case-control study of adult malignant brain cancer. *Cancer Epidemiology, Biomarkers and Prevention*, *17*(10), 2639–2646.
- Link, M. W., & Mokdad, A. H. (2005). Alternative modes for health surveillance surveys: an experiment with web, mail, and telephone. *Epidemiology*, *16*(5), 701–704.
- Link, M. W., & Mokdad, A. H. (2005). Effects of survey mode on self-reports of adult alcohol consumption: A comparison of mail, web and telephone approaches. *Journal of Studies on Alcohol*, *66*(2), 239–245.
- Harewood, G. C., Yacavone, R. F., Locke, G. R, 3rd, & Wiersma, M. J. (2001). Prospective comparison of endoscopy patient satisfaction surveys: E-mail versus standard mail versus telephone. *American Journal of Gastroenterology*, *96*(12), 3312–3317.
- Hirschfeld, S., Songco, D., Kramer, B. S., & Guttmacher, A. E. (2011). National Children’s Study: Update in 2010. *Mount Sinai Journal Medicine*, *78*(1), 119–125.
- Williams, L. M., Morrow, B., Lansky, A., et al. (2003). Surveillance for selected maternal behaviors and experiences before, during, and after pregnancy. pregnancy risk assessment monitoring system (PRAMS), 2000. *MMWR Surveillance Summaries*, *52*(11), 1–14.
- Sinclair, M., O’Toole, J., Malawaraarachchi, M., & Leder, K. (2012). Comparison of response rates and cost-effectiveness for a community-based survey: Postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Medical Research Methodology*, *12*, 132.