Biological specimens for community-based surveillance studies: Method of recruitment matters

Brenda L. Coleman
University of Toronto

Iris A. Gutmanis
St. Joseph’s Health Care, London

Susan J. Bondy
University of Toronto

Allison J. McGeer
Mount Sinai Hospital, Toronto

Marina I. Salvadori
Children’s Hospital of Western Ontario, London

Marie Louie
Provincial Laboratory of Public Health, Calgary

Studies requiring the collection of biological specimens are often difficult to perform and costly. We compare face-to-face and telephone interviews to determine which is more effective for return of self-collected rectal swabs from subjects living in rural and semi-rural areas of Ontario, Canada. People interviewed face-to-face in 2006-2007 were asked to provide a rectal swab while the interviewer waited. Those interviewed by telephone were sent a package and asked to return the swab by mail, with one follow-up reminder call. Telephone interviewing resulted in a higher response rate for the completion of household and individual-level questionnaires. However, face-to-face interviews resulted in a significantly higher proportion of interviewees who returned swabs making the participation rate higher for this mode of contact (33.7 versus 25.0 percent). Using multivariable logistic regression, higher rates of rectal swab return were associated with face-to-face interviewing while adjusting for the impact of household size and respondent age and sex. For studies requiring the submission of intimate biological samples, face-to-face interviews can be expected to provide a higher rate of return than telephone interviews.

Keywords: biological specimens, data collection, health surveys, interviews

1 Introduction

Large scale population-based studies and comprehensive case-control studies requiring participant interview data and the collection of biological specimens are expensive and require exhaustive coordination efforts. Non-participation may affect estimates of outcomes and impinge upon the generalizability of the research findings (Groves 2006). Further, low participation rates are costly in terms of the time and money spent on the identification of eligible participants and initial contact.

The recruitment of community-dwelling subjects is an ongoing challenge for researchers who must select a data collection mode that meets the study requirements and is also acceptable to study participants. Face-to-face interviews cost substantially more than those done by telephone or mail (Edwards et al. 2002; Weeks et al. 1983) yet may be necessary when anthropological measurements or biological specimens are needed. On the other hand, recruitment by telephone has been made increasingly difficult by reliance on answering machines or voicemail, call display, and cellular phones as well as the public’s frustration with telemarketers (Curtin et al. 2005; Kempf and Remington 2007; Scheaffer et al. 1996). For specimens that can be collected by the participant, there is a choice of strategy including face-to-face, post, internet, or telephone, but a paucity of information available on participation rates by mode of contact. There is even less information regarding participation rates for studies requiring the collection of samples that may be deemed sensitive by the participant.

The Ontario Well Water Study required the collection of rectal swabs for detection of antimicrobial resistant bacteria. Two previous postal surveys that required mailing of a stool sample to determine the prevalence of antimicrobial resistant Escherichia coli in European and Canadian community-dwelling subjects experienced low participation rates at 24 and 26 percent, respectively (Bruinsma et al. 2002; Bruinsma et al. 2003). Other studies that asked subjects to send self-collected biological samples following a telephone interview or mailed questionnaire had response rates that varied from 13 to 74% (Bauer et al. 2004; Cozier et al. 2004; Domeika et al. 2007; Engel et al. 2002; Macleod et al. 2005; Olshan et al. 2007; Rogstad et al. 2001). In comparison, somewhat higher rates of return of biological specimens were reported (70 to 96%) following face-to-face interviews (Akwar et al. 2007; Fenton et al. 2001; Klavs et al. 2004; Klavs et al. 2002; McCadden et al. 2005; Turner et al. 2002). Study investigators determined that face-to-face interviews may re-
sult in a higher rate of rectal swab return and this strategy was implemented among a non-randomized subset of targeted households. The remainder of household contacts were interviewed by telephone with postal return of specimens. This analysis examines the effect of data collection mode on response rates and rates of return of rectal swabs for subjects in one Canadian province.

2 Methods

The present study was a cross-sectional study of factors associated with the human carriage of antimicrobial resistant *Escherichia coli* nested within a case-control study of households that submitted water samples for bacteriological laboratory analysis at seven public health laboratories in Ontario, Canada. In Canada, there is no registry of private drinking water wells. Private well water testing is voluntary but universally accessible through regional laboratories. All *E. coli* positive water samples submitted for testing between May 1, 2005 and September 30, 2006 to two western regional laboratories (London and Hamilton) as well as a randomly selected monthly quota of samples from five eastern regional laboratories (Ottawa, Kingston, Peterborough, Orillia, and Toronto) were screened for susceptibility to antimicrobial agents to determine the prevalence of antimicrobial resistant *E. coli* in private water sources (Mataseje et al. 2009). A study investigating risk factors associated with contamination of private water sources was conducted using these water samples to identify cases and controls (Coleman et al. 2009). For the case control study, cases were households with water samples that tested positive for antimicrobial resistant *E. coli*. Controls were randomly selected households with water samples yielding *E. coli* susceptible to antibiotics. A second set of controls consisted of households randomly selected from water samples with no bacterial contamination. Inclusion in the case-control study was limited to unique households that provided samples from private drinking water sources that also provided an operational telephone number with the water sample and had at least one adult (18 years or older) resident who spoke English well enough to participate and who consented to share their contact information with the study.

Households participating in the case-control study were asked to participate in the cross-sectional study in which all household members 12 years and older were eligible for inclusion.

During the study period, all households that submitted a water sample were mailed a study information sheet along with the results of their water test. Then a study assistant at the Safe Water Unit of the Ontario Ministry of Health and Long-Term Care telephoned selected households and, using a script in which they identified themselves as being from the Safe Water Unit, gave a very brief description of the study and the study investigators (doctors and scientists from the universities) and asked if the primary contact was willing to share their name and phone number with the study group. Information from eligible and consenting households was forwarded to the appropriate interviewer.

The third exchange with the household varied by mode of contact. In the western laboratory regions, 35 percent of households were recruited by telephone with appointments set to visit the subject in their home. To reduce the number of people making contact with the household as much as possible, the site interviewers made the recruiting/appointment calls themselves. Upon contact with an adult in the household, they identified themselves as working with researchers at the associated universities and Ministry of Health and Long-Term Care, gave a brief explanation of the study, and asked if the household contact would be willing to participate. If the contact agreed, the interviewer set up a time to visit the home, asked permission to speak with other eligible household members during the scheduled visit, and provided their first name and coordinator’s contact information for anyone needing to reschedule. The interviewer called to confirm the day, time, and location of the visit 24 to 48 hours before the scheduled appointment.

During the site visit, interviewers explained the objective of the study and its requirements. After obtaining written consent, a household member was asked to complete a dwelling questionnaire regarding factors that might affect drinking water quality. Personal questionnaires were then completed only for eligible household members who agreed to submit a rectal swab. Following completion of the personal questionnaire, interviewers explained how to collect the swab, provided the participant with a collection kit including swab collection instructions, swab, transport media, pre-numbered label, and biohazard bag, and asked them to collect the swab during the site visit. Only one visit was made to each household.

Household members in the remaining 65 percent of households in the western regions and 100 percent of households in the eastern regions were contacted and interviewed by telephone. Trained interviewers at a reputable, private-sector survey company called the household, explained the study, and collected data from consenting household contacts for the dwelling questionnaire. They also asked permission to speak with other eligible household members. If they were not available, they asked for a time to make a return call to speak with them. Personal interviews were completed only for household members who agreed to submit a swab. Rectal swab sampling kits were subsequently mailed to the participants. If more than one household member agreed to collect a swab, the kits were posted directly to each participant to increase the perception of personal contact as well as to avoid confusion. The mail out package included a letter of introduction, swab collection kit (described above), stamped self-addressed envelopes, and consent forms. A follow-up telephone call was made to participants who had not returned the rectal swab within three weeks of mail out. A second kit was mailed to subjects who did not receive the original kit.

In all cases, a minimum of ten attempts, at different times of day and days of the week, were made to reach the household at each of the telephone contact points. Written consent was required of all participants. Parental permission was established before talking with youths 12 to 18 years of age. For youths 12 to 15 years old, verbal assent was required.
followed by written parental consent. Youths 16 to 18 were required to provide written consent. Everyone interviewed by telephone was required to give verbal consent prior to the interview and to provide written consent with rectal swab submission. Study ethical approvals were received from associated universities and no remuneration or other incentives were offered to participants. Households were informed of whether or not their water sample was contaminated with E. coli as per standard procedure. However, they were not informed of the results of the antimicrobial resistance screening.

As stated, no population registry of private wells exists in Ontario and no data on non-consenting households were available from the Ontario Ministry of Health and Long-Term Care. Thus, non-response bias is assessed using the 2006 census data (Statistics Canada 2008) of census divisions serviced by the seven participating regional public health laboratories. As the overwhelming majority of private wells are located in rural areas, the largely urban census metropolitan areas within the census divisions are not included in the estimates. A range of household incomes for study participants used the mid-points for each income categories except the highest category ($80,000 and over) for which we used $90,000 and $100,000, respectively, to estimate the stated range.

Standard estimates using the American Association of Public Opinion Research definition 3 (American Association for Public Opinion Research 2006) are presented for the initial contact response rate. Subsequent rates of response are direct calculations of participation and refusal. The estimated net probability of receipt of a rectal swab conditional on mode of interview is calculated using the probability of participation at each preceding point of contact. All p values are based on two-sided tests of significance and all confidence intervals are based on a type I error of five percent.

Variables associated with return of a rectal swab (participation) were assessed using logistic regression (1 returned swab; 0: no swab) adjusted for clustering within households using the Taylor linearized variance estimation obtained through the complex survey commands in Stata version 9.2 (2007). The primary sampling unit was the household and since no selection of participants within households occurred (i.e. all members 12 years and older were eligible to participate), no sampling weights were used. Following bivariate analysis of variables, those associated with returning a swab at a p value of <0.25 were selected for inclusion in the full model. To obtain a parsimonious model, variables were removed from the model according to a modified form of the strategy for a focal hypothesis as suggested by Vittinghoff (Vittinghoff et al. 2005). Covariates and interaction terms were retained if they were significantly associated with the outcome (through Wald tests) or if their retention modified the observed association between interview mode and the outcome (i.e., an observed confounding effect).

Continuous variables were kept in original form after confirming that the assumption of linear associations was met. Residual diagnostics did not detect any outlying observations and important multicollinearity was not present as assessed through variance inflation factor statistics. Model-related odds ratios and confidence intervals for variables involved in interactions are presented for linear combinations of coefficients.

3 Results

3.1 Household level participation

Of the 342,009 water samples submitted for testing during the study period, 2,351 were selected for the study, of which 1,771 were eligible. A response rate of 68.6% was achieved with 1,190 of the 1,771 households agreeing to share their name and telephone number with the study see Figure 1. There was no difference in household response rates by laboratory region (68.6 and 69.4% for the western and eastern regions, respectively) or by laboratory result (68.2 versus 69.3% for E. coli contaminated water versus no contamination).

A dwelling questionnaire was completed for 879 (49.6%) eligible households. As shown in Table 1, households that completed dwelling questionnaires were similar in education level, household income, and results of their water test by mode of contact. However, a higher proportion of households interviewed face-to-face (33%) than by telephone (23%) were located on a farm (p=0.003) and dwelling questionnaires were more likely to be completed by telephone than if a site visit was required (55 versus 42%; p <0.001). The average size of participating households (2.8; 95% CI: 2.7, 2.9) and the average household income ($67,712-71,959) were similar to the 2006 Canadian census of the non-urban areas included in the study regions at 2.7 and $70,725, respectively.

3.2 Individual level participation

Personal interviews were completed by people who agreed to subsequently submit rectal swabs. Interviews were completed by 1,038 of 2,101 (49.4%) eligible people within the participating households. While over half of adults (54.5%) within participating households completed a questionnaire, only 9.3% of adolescents, 12 to 19 years old, were completed by 1,038 of 2,101 (49.4%) eligible people within the participating households. While over half of adults (54.5%) within participating households completed a questionnaire, only 9.3% of adolescents, 12 to 19 years old, completed one (p <0.001). Personal questionnaires were equally likely to be completed by females (n=525) and males (n=513).

The probability of completing a personal questionnaire was higher for people interviewed face-to-face (57.7%) than by telephone (46.8%; p <0.001). This difference was driven by the adult respondents with 63.7% of adults in households interviewed face-to-face completing a personal interview compared to 51.6% of those interviewed by telephone (p <0.001). Conversely, there were equal proportions of eligible adolescents in each mode of contact (9.4 and 9.3%, respectively; p=1.00) who completed a personal questionnaire.

3.3 Individual level participation: Return of rectal swab

Swabs were submitted by 784 of the 1,038 subjects (75.5%) who agreed to submit one prior to completing a per-
sonal questionnaire. The probability of submitting a rectal swab, conditional on mode of contact, was higher for those interviewed face-to-face (33.7%) than for those interviewed by telephone (25.0%; \( p < 0.001 \)).

Bivariate analysis of household and personal variables revealed that the odds of returning a swab was higher for people interviewed face-to-face, for older respondents, and for people from households with fewer eligible respondents. There was no statistically significant association with household income, household education level, property type, or whether or not the water was contaminated with *E. coli* (Table 2).

The final logistic regression model, which adjusts the estimates for the effects of other variables in the model (Table 3), indicates that the odds of returning a swab were 1.3 times higher for every decade increase in age and 1.2 times higher per person decrease in the number of eligible respondents within the household. A statistically significant interaction was found between the sex of the respondent and the mode of interview in relation to the probability of swab return. Participants were generally more likely to return a swab following a face-to-face than telephone interview with the odds of return of a swab higher for females (46.0; 95% confidence interval (CI): 10.9, 193) than males (5.7; CI: 2.8, 11.5) following the completion of a face-to-face compared to a telephone interview.

### 4 Discussion

Our analysis found that rates of participation in a population-based study requiring provision of a self-collected rectal swab varied both by the stage of participation and, for specimen collection, whether the interview was done face-to-face in the respondent’s home or by telephone. We found that completion of the household interview was higher by telephone than for face-to-face interviews. However, the return of swabs was so much higher for those interviewed face-to-face relative to those interviewed by telephone that this more than compensated for the difference: the net probability of returning a rectal swab favoured the use of face-to-face specimen collection at 33.7% versus 25.0% for return of the specimen by post. While low in either condition, these participation and submission rates are consistent with other studies in this area of research in which stool samples or rectal swabs were required and in which no remuneration or other incentives were provided (Bruinsma, Filius et al. 2003; Bruinsma, Hutchinson et al. 2003).

We acknowledge that difference in rates of participation in the household interview may be partly influenced by interviewer experience. Interviewers in the telephone data collection condition were staff of a professional survey centre while university students were employed for the face-to-face interviews. However, all interviewers were research trained and the recruitment scripts were identical. Also, people in the face-to-face group were required to commit to a home visit which likely reduced participation.
Methodological literature for surveys that require biological samples is limited relative to that for self-report data. The comparison of rates of return of biological samples is further hampered due to the variety of study populations, study methods, the nature of the specimen being obtained, and because there are few publications directly comparing modes of collection for community-dwelling subjects. One study completed in 1995 in Nova Scotia, Canada found that, similar to our findings, face-to-face recruitment of adults resulted in higher questionnaire response and attendance at clinics for collection of blood samples (36.9%) than recruitment by post (31.5%) (Eastwood et al. 1996).

As seen in our study, even though people participate in one part of the study by completing an interview, and although they agree to provide a self-collected biological specimen, only 22-74% of people interviewed by telephone (Cozier et al. 2004; Domeika et al. 2007; Engel et al. 2002; Macleod et al. 2005) and 66-83% interviewed face-to-face actually returned/provided specimens (Klavs et al. 2004; Klavs et al. 2002; McCadden et al. 2005; Turner et al. 2002). This difference may be attributable to elevated motivation to participate when the interviewer is face-to-face with the subject (Fenton et al. 2001; Frey and Oishi 1995). The presence of the interviewer in face-to-face surveys may increase the probability of participation due to the participant’s perception of the importance and legitimacy of the study and a higher level of trust engendered through personal interaction (Robling et al. 2010). It is also much easier to passively decline, by not completing all study requirements, when an interviewer is not physically present (Stoop 2005).

Unlike a number of other studies with biological sample requirements (Goldstein and Jennings 2002; Jackson et al. 1996; Kozlowski et al. 2002; Mishra et al. 1993; Olshan et al. 2007; Søgård et al. 2004), we found no difference in the proportion of people who returned a rectal swab by household income and, although there was a trend towards higher participation with higher levels of education, the effect was not statistically significant. However, we did find that rates of return of swabs differed by the respondent’s age. For respondents interviewed by telephone, rates of swab return generally increased with age. This mirrors findings of other studies involving community-dwelling subjects asked to submit biological samples following telephone or postal surveys. Increasing age was associated with higher rates of
Table 2: Bivariate associations between return of rectal swab and covariates. Logistic regression adjusted for household clustering, Ontario Well Water Study, 2006-2007

<table>
<thead>
<tr>
<th>Variable (range)</th>
<th>Odds ratio</th>
<th>95% C.I.</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent’s age (12-89)</td>
<td>1.03</td>
<td>1.01, 1.04</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mode of contact</td>
<td>Telephone</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>Face-to-face/site visit</td>
<td>11.1</td>
<td>5.79, 21.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of eligible people in household (1-10)</td>
<td>0.80</td>
<td>0.69, 0.94</td>
<td>0.006</td>
</tr>
<tr>
<td>Respondent’s sex</td>
<td>Female</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>Male</td>
<td>1.25</td>
<td>0.98, 1.58</td>
<td>0.07</td>
</tr>
<tr>
<td>Water test result</td>
<td>Not contaminated</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>Contaminated with <em>E. coli</em></td>
<td>0.78</td>
<td>0.54, 1.10</td>
<td>0.16</td>
</tr>
<tr>
<td>Household income (2005)</td>
<td>&lt;$40,000</td>
<td>0.90</td>
<td>0.75, 1.07</td>
</tr>
<tr>
<td>$40,000-59,999</td>
<td>2.73</td>
<td>1.38, 5.41</td>
<td>0.004</td>
</tr>
<tr>
<td>$60,000-79,999</td>
<td>0.93</td>
<td>0.50, 1.73</td>
<td>0.83</td>
</tr>
<tr>
<td>$80,000 or more</td>
<td>0.86</td>
<td>0.52, 1.42</td>
<td>0.55</td>
</tr>
<tr>
<td>Not stated</td>
<td>0.79</td>
<td>0.49, 1.26</td>
<td>0.32</td>
</tr>
<tr>
<td>Household education</td>
<td>Less than secondary grad</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>Secondary school grad</td>
<td>0.97</td>
<td>0.47, 2.01</td>
<td>0.93</td>
</tr>
<tr>
<td>College or trade school</td>
<td>1.17</td>
<td>0.61, 2.27</td>
<td>0.63</td>
</tr>
<tr>
<td>University</td>
<td>1.43</td>
<td>0.75, 2.72</td>
<td>0.28</td>
</tr>
<tr>
<td>Not stated</td>
<td>0.92</td>
<td>0.37, 2.29</td>
<td>0.85</td>
</tr>
<tr>
<td>Property type</td>
<td>Farm</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>Non-farm</td>
<td>1.08</td>
<td>0.72, 1.61</td>
<td>0.72</td>
</tr>
</tbody>
</table>

* Robust variance estimate
** Wald chi square test
† Highest education level in household
‡ Probability that at least one of the variables’ regression coefficients ≠ 0

return from community-dwelling respondents asked to provide buccal swabs for DNA analysis (Kozlowski et al. 2002), semen samples (Olshan et al. 2007), or blood samples (Chan et al. 2007; Eastwood et al. 1996; Hará et al. 2010; Macleod et al. 2005; Søgarrd et al. 2004). Our rate of swab return peaked around 60 years of age and then dropped off again for older adults which is similar to patterns noted by some other researchers (Boshuizen et al. 2005; Keinan Boker et al. 2001; Lee et al. 2008; Ronckers et al. 2004). For studies requiring biological samples from a wide range of ages, sample size estimation and study methods need to take into account potential differences in participation rates by age group.

We also found that, even though an equal number of males and females completed interviews and verbally agreed to submit a swab, females were modestly more likely than males to submit one following a face-to-face interview while males were more likely to post a sample following a telephone interview. The reason for this difference, which was largely driven by the difference between males and females 70 years of age and older, is not clearly understood and requires further study. We also found that the odds of returning a swab were higher for people in households with fewer eligible respondents. This finding was not dependent on the sex or ages of the people within the household. There are myriad factors that can affect the propensity to participate in research on different populations and subject areas. The interaction of social, personal, interviewer, and survey-level factors is difficult to parse and additional research is needed to understand how best to address them and the impact of nonparticipation on survey estimates, whether they are self-report or biological measures (Groves 2006; Groves and Peytcheva 2008).

There are several limitations to our study. The main goal of the research project was to complete the case-control and cross-sectional studies described in the methods section. Thus, the households were not randomly assigned to each mode. Also, our study was limited to a convenience sample of households that submitted private water samples for bacteriological testing. Although we found no difference in response based on whether the water source was contaminated with *E. coli* or not, generalizability may be limited. Another limit to generalizability is that households in this study were located in rural areas of the province so the findings may not apply to urban dwelling populations. However, only 25% of participants lived on a farming property and many Canadians living in rural areas, including those living on farming properties, are employed in urban areas (Alasia and Bollman 2009).
For studies requiring the submission of biological samples, face-to-face interviews can be expected to provide a higher rate of return than telephone interviews and postal submission of specimens, especially for adult respondents at either extreme of the age range. However, the trade-off in cost per returned sample may only be justifiable in studies in which the enrolment of subjects is challenged by the rarity of the condition or the cost of recruiting each subject.

Future studies using different methods of recruitment and inducements for participation are required to determine if the overall response rates for community-based surveys requiring submission of biological samples, especially ones that may be considered sensitive or embarrassing to collect or share, can be improved.

References


Fenton, K. A., Copas, A., Mitchell, K., Elam, G., Carder, C., Ridg-


