OBTAINING DEMOGRAPHICALLY RELEVANT DATA THROUGH CELL PHONES:
LESSONS FROM A PILOT STUDY IN CEBU

by

Socorro A. Gultiano, Michelle J. Hindin, Mitali Thakor

For additional information, please contact:

Authors’ name       Socorro A. Gultiano*, Michelle J. Hindin, Mitali Thakor
Designation         Senior Research Associate*, Associate Professor, Researcher
Affiliation          University of San Carlos Office of Population Studies
                      Foundation*, John Hopkins University, John Hopkins University
Address              Office of Population Studies, University of San Carlos,
                      Talamban, Cebu City*
Tel. no.              +6332-3460102, +6332-3466050
E-mail               conniegultiano@gmail.com
OBTAINING DEMOGRAPHICALLY RELEVANT DATA THROUGH CELL PHONES: LESSONS FROM A PILOT STUDY IN CEBU

by

Socorro Gultiano, Michelle Hindin and Mitali Thakor

ABSTRACT

In the Philippines, information on sexual activity and contraceptive use has, in most cases, been collected by means of face-to-face (FTF) interview. This has been the case notwithstanding researchers’ awareness of the vulnerability of FTF to social desirability bias particularly with questions related to sensitive behavior. In a few studies, self-administered questionnaires (SAQ) have been resorted to, but little has been written about the pros and cons of this method vis-à-vis FTF in the Philippine setting.

This study introduces another alternative mode of data collection - mobile phone-based interviews using the Episurveyor software. The data are from a pilot study of 32 young adult participants (16 males, 16 females) from the Cebu Longitudinal Health and Nutrition Survey. Cell phones were programmed and distributed to the participants who were then asked to respond to questions in their cell phones about sexual activity and contraceptive use on a day-to-day basis over a four-week period (also referred to as “experience sampling method” or ESM). The study design also included a baseline FTF, an exit phone and FTF interview and SAQ.

The data will be analyzed with respect to: 1) the acceptability of the cell phone mode of data collection in comparison to FTF and SAQ, 2) completion and response rates for the three methods of data collection, and 3) comparison of the results on sexual activity and contraceptive use of the cell phone-based method vis-à-vis those of FTF and SAQ (correlations). Other pragmatic concerns arising from the study will also be discussed.

Introduction

This study introduces a novel method of collecting demographic and health-related data in a developing country setting such as the Philippines. Cell phones are used to collect data on sexual activity, contraceptive use, and mood in a pilot study of young adults in Metro Cebu using the Experience Sampling Method (ESM). ESM is a structured diary method of data collection that captures an individual’s daily life experiences and is ideal for obtaining information about location, context, activities and feelings [1]. ESM methods signal participants about when to respond to questions from the researcher. The use of mobile phone technology is therefore suitable for ESM. The popularity and pervasive use of cell phones in the country provide additional motivation for trying out mobile phone-based data collection as an alternative to the conventional paper-based interviews. This alternative becomes particularly useful when the information to be collected is deeply personal and sensitive, such as day-to-day sexual activity, contraceptive use and feelings.

The rapid development of mobile technologies in the past decade has opened new avenues for research, particularly in developing countries where mobile technologies have “leapfrogged” ahead of developed regions in terms of near-ubiquitous penetration. For example, today there are approximately 5 billion mobile devices in use around the world, and close to two-thirds of them are in the hands of people living in emerging market economies.

1 Socorro Gultiano is a Senior Research Associate of the University of San Carlos Office of Population Studies Foundation; Michelle Hindin is an Associate Professor of the Department of Population, Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health; and Mitali Thakor is a graduate student at the Massachusetts Institute of Technology.
Numerous examples exist of technologies working toward improving medical practitioner connectivity, patient care, and collection of community and clinical health data [2-5]. However, while many ongoing public health projects address mobile-based interventions, very little research exists particularly on understanding feasibility and success of mobile-based data collection. Electronic methods of data collection, such as audio-enhanced, computer-assisted self-interviewing (ACASI) have been increasingly used instead of paper-based methods throughout the past decade [6-8]. It is presumed that more confidential methods of collecting data would reduce reporting bias, however results from studies that have used ACASI have yielded mixed results [9-11]. More recently, researchers have also begun to experiment with mobile electronic devices such as handheld personal digital assistants (PDAs) for collecting data in resource-constrained settings [10-11].

In the Philippines, cell phones present an attractive alternative to ACASI and PDAs for various reasons including: 1) reduced cost of the device, 2) a wider reach of wireless networks over fixed line internet access, and most importantly, 3) the increased likelihood of participant familiarity with the device, which helps reduce training time and costs. Recent studies on mobile phone data collection have mentioned other advantages of cell phones over PDAs, such as the preference of adolescents and young adults for their use [12], real-time detection of research-adverse events during data collection [13], and reduced data loss and real-time supervision of interviewers and community health workers [14].
Objectives

This study is interested in assessing the feasibility, practicability, and accuracy of cell phone data collection for demographically relevant but sensitive data such as sexual activity and contraceptive use. Cell phone-based ESM is well suited for obtaining information on coital frequency, the least frequently measured and possibly most underutilized determinant of fertility. Understanding the frequency of sexual activity also has important reproductive health implications. In the Philippines, nearly all surveys on sexual behavior and reproductive health have been done with face-to-face interviews, despite growing evidence from other developing countries that shows that other survey modalities may elicit more responses about sensitive behaviors and minimize social desirability bias [15-18]. Electronic surveys to report sensitive behaviors may be an alternative for young adults in a developing country that is technologically advanced, particularly with mobile technologies: the Philippines has seven times the number of mobile phones compared to landlines, and Filipinos send over 200 million cell phone text messages a day, ten times the per capita world average [19].

While cell phone-based data collection purportedly offers a number of advantages over paper-based interviews when collecting sensitive data, these advantages need to be weighed against the costs of technology, interviewer capabilities for training and trouble shooting, field time, as well as respondent burden. While survey staff will have less data entry time with methods like cell phone-based ESM, the programming and learning curve for staff and participants can be steep. Exploration of these pros and cons of the technology is what prompted this pilot study. Results of this study will help identify strengths and weaknesses and recommend adjustments where needed for larger-scale studies on ESM and other related surveys.

Data and Methods

Study Setting and Sampling

The data reported here are part of the Cebu Longitudinal Health and Nutrition Survey (CLHNS), an ongoing data collection project of the USC Office of Population Studies (OPS) in collaboration with the UNC Carolina Population Center, Johns Hopkins University and other American universities. The CLHNS followed up all pregnant women in 33 randomly selected communities in Metro Cebu during their 3rd trimester of pregnancy, and who subsequently gave birth in May 1983–April 1984. Follow-up has continued for these women and the birth cohort in subsequent years: 1991, 1994, 1998, 2000, 2002, 2005 and 2008. The data presented here are based on the pilot study of 32 members of the birth cohort (16 males and 16 females) who were asked to complete daily diaries on cell phone for a one-month period. The pilot study pre-selected men and women who owned cell phones (to minimize possibility of theft/loss of the device and ensure familiarity of use) and who had no intention of moving out of Metro Cebu during the study period. Pregnant and lactating women were excluded from the study since pregnancy and lactation may affect sexual and contraceptive behavior. The study participants were 25-26 years old at the time of the study.

Study Design

Experienced interviewers of OPS, four males and four females, collected the data for this study. They were proficient in cell phone usage and were trained for one week on the data collection software (Episurveyor) and troubleshooting. All pre-selected participants of the study were screened prior to the study to determine eligibility based on marital/cohabiting status (half of the sample were to be never-married and the other half, ever-married,
including cohabitation), reproductive status (pregnant and lactating women were to be excluded), and cell phone ownership.

All participants were asked for their consent to participate in the study and pledge to return the cell phone in good condition at the end of the study. If they consented, they were given a face-to-face (FTF) baseline survey after which they were given training on the use of a cell phone pre-installed with the survey software and prompt via an alarm or SMS (sent from the staff phone) to fill out the survey every day for four weeks. The daily surveys began with questions establishing respondents' location and status, followed by a series of questions on sexual thought and activity, and contraceptive use if sexual activity had taken place. The final series of questions dealt with mood and emotions, a topic not covered in this paper. After four weeks, the participants were asked to complete an exit survey on the phone as well as a FTF interview (for half of the sample) and a paper-based self-administered questionnaire (SAQ) (for the other half). The exit survey asked about sex, contraception and mood over the past four weeks, and their preferences and thoughts about using the cell phones for the survey.

Because respondent burden is heavier in cell phone-based ESM studies than in conventional paper-based interviews, the proponents of the study felt that it was necessary to provide incentives to the participants. A scheme unique to the cell phone focus of the study was implemented. Participants were compensated in increments of cell phone airtime (cell phone “loads” for their own cell phones) at weekly intervals throughout the study: P100 worth of “load” at baseline, P200 at the completion of each of the 3 succeeding weeks of surveys, and P100 at exit interview. This compensation was regarded also as a deterrent to theft and careless handling of the project phone by the participants.

**EpiSurveyor Mobile Data Collection Software**

The study used Java-enabled Nokia 3120c phones installed with the EpiSurveyor software, a free data collection application developed by DataDyne, Inc. Cell phones were purchased locally at the unit price of P5,185, and were determined to be middle-range, fairly common devices at the time of the study. The EpiSurveyor platform was selected after piloting several mobile data collection tools (e.g., FrontlineSMS), based on its specific design for public health data collection and availability in free and open-source format. EpiSurveyor was downloaded to the cell phones via GPRS prior to the survey. Password locks were installed on the phone and application. Survey completion took place offline, and the surveys were saved to the phone SIM immediately after completion. The project staff visited the participants every week to copy files from SIM to external memory cards to reduce risk of data loss. Admittedly, the most convenient and secure method of saving the data would have been automatic uploading to the DataDyne server via GPRS, however internet access was intermittent in many of the project sites and, also, it seemed prudent to disable participants’ use of web applications on the study devices. At the completion of the study, survey data were exported into standard file formats (CSV and Microsoft Excel) to enable review and analysis.

**Data Analysis**

We focus on the responses to questions on sexual activity and contraceptive use of 31 respondents who participated in the 4-week study. One male respondent had dropped out of the study after baseline interview because the cell phone given to him got broken when his 5-year old son allegedly played with it and dropped it.

We examine survey completion of the ESM data, the acceptability of being asked about sexual activity on cell phones, preferences for using phone surveys to collect data on sexual activity, and the agreement between responses on cell phones versus other modes of
reporting sexual activity such as paper-based interviews through FTF and SAQ. Where needed, we present results stratified by gender and marital status to reflect the design of the pilot study.

Results

Sample Characteristics

Of the 16 male participants in the study, 8 were single and 8 were ever-married. Among the 16 female participants, 7 were single and 9 were ever-married (in the baseline interview it was found that the pre-selected “single” female was actually “separated”). As mentioned earlier, one married male participant had dropped out of the study leaving only 31 participants with ESM and exit data for analysis.

On the average, the respondents had 10.3 years of formal education (std dev=2.0). About 40.6% of the sample was working at the time of the survey, with more married respondents working (58.8%) than their single counterparts (20.0%).

Completion of ESM data

In total, we collected 815 person-days of diary data with an average of 26.3 completed survey days per respondent (std dev= 7.1). No significant differences by gender or marital status was observed in terms of survey days completed.

Although the mean number of survey days completed by the participants is relatively high (with maximum days of participation being 28 or 29 days), it was observed that there were a number of cases showing duplicate or missing responses on a day-to-day basis. Duplicates meant that the respondent answered and saved the survey twice or more in one day, while missing cases meant that the respondent failed to answer or save the survey on that day. In all, 25.8% of the respondents correctly completed (without misses and duplicates) all survey days, with more single than ever-married respondents correctly completing the survey (33.3% vs. 18.7%); 19.3% had duplicates, 25.8% had misses, while 29.0% had both misses and duplicates. In one case, a female respondent had completed only 7 days of survey because her American fiancé took her to Manila to process her visa and was angry when he found out about the survey, prohibiting her to continue answering any further. Another female respondent had as many as 48 dairy data because she failed to turn off the daily alarm that prompted her to respond to the survey (she just set it to “snooze”) and thus she would answer the survey again when the alarm went on again. Obviously a host of factors contributed to the missing and duplicate dairy data, some of which could be (and had been) mitigated by more intense training and troubleshooting, while others were simply beyond our control.

Preferences and Acceptability

Overall, 93.6% of the participants reported that they would rather be asked about sexual activity on the phone than to be interviewed. This pronounced preference for phone survey prevailed regardless of gender, marital status, and status of survey completion (i.e., whether ESM was complete, had duplicates or misses). There were no differences in preference by whether the exit interview was done by an interviewer (n=15) or by SAQ (n=16).

With respect to accuracy, 87% of all participants said that they answered questions about sexual activity more honestly on the phone than in the paper interview. Among those whose exit interview was with an interviewer, 80.0% reported answering more honestly about sexual activity on the phone than in the interview, whereas 93.7% said they answered more
honestly on the phone than on the SAQ. When asked why they would answer more honestly on the phone than on paper interview, 48.4% of the respondents mentioned more privacy with the phone, 25.8% mentioned being less embarrassed on the phone, and 22.6% said that it was faster to answer on the phone.

Measuring Sexual Activity

Over the 4-week period that ESM data was collected via cell phone, 48.4% of the sample reported no sex during the period of observation (40.0% of males and 56.2% of females; 73.3% of the never married and 25.0% of the ever married). Of the 9 males and 7 females (total of 51.6%) who did report having sex, 8 and 6 of them, respectively, reported having used some means of contraception. All 4 of the never-married respondents also reported using a family planning method, while 10 out of the 12 ever-married respondents reported having done the same.

If we focus on the week preceding the baseline and exit surveys, we see that only 29.0% of the respondents reported having had sex before the baseline interview compared with 35.5% reporting having had sex on the week before the exit survey. The ESM data is even more interesting since the proportion reporting sex on the last week of the study (which roughly coincides with the week prior to exit interview) is much higher at 45.2%. Except for female respondents, the proportion reporting sex during the preceding week (among males and among never-married and ever-married respondents) is persistently higher with the cell phone-based ESM data.

As far contraceptive use is concerned, it is observed that the proportion reporting use is lower than the proportion that reported having had sex, with the exception of male respondents who all reported using contraception if they reported having had sex in the last week. Overall, report of contraceptive use was also highest in the ESM data (38.7%) than in the paper interviews (29.0% in the exit survey and 22.6% in the baseline survey).

Table 1. Reporting of sexual activity and contraceptive use, by type of survey, gender and marital status

<table>
<thead>
<tr>
<th></th>
<th>Male (n=15)</th>
<th>Female (n=16)</th>
<th>Never married (n=15)</th>
<th>Ever married (n=16)</th>
<th>Total (N=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Had sex in the last week</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>5 (33.3%)</td>
<td>4 (25.0%)</td>
<td>1 (6.7%)</td>
<td>8 (50.0%)</td>
<td>9 (29.0%)</td>
</tr>
<tr>
<td>ESM</td>
<td>7 (46.7%)</td>
<td>7 (43.8%)</td>
<td>4 (26.7%)</td>
<td>10 (62.5%)</td>
<td>14 (45.2%)</td>
</tr>
<tr>
<td>Exit</td>
<td>4 (26.7%)</td>
<td>7 (43.8%)</td>
<td>2 (13.3%)</td>
<td>9 (56.2%)</td>
<td>11 (35.5%)</td>
</tr>
<tr>
<td><strong>Used contraceptives in the last week</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>5 (33.3%)</td>
<td>2 (12.5%)</td>
<td>0 (0.0%)</td>
<td>7 (43.8%)</td>
<td>7 (22.6%)</td>
</tr>
<tr>
<td>ESM</td>
<td>7 (46.7%)</td>
<td>5 (31.2%)</td>
<td>4 (26.7%)</td>
<td>8 (50.0%)</td>
<td>12 (38.7%)</td>
</tr>
<tr>
<td>Exit</td>
<td>4 (26.7%)</td>
<td>5 (31.2%)</td>
<td>1 (6.7%)</td>
<td>8 (50.0%)</td>
<td>9 (29.0%)</td>
</tr>
</tbody>
</table>
If degree of concordance is examined with respect to responses on sexual activity on the week preceding the surveys, it is intriguing to note that, overall, type of survey (mode of data collection) figures just as prominently as reference period in establishing congruence of responses. As demonstrated in Table 2, responses in the baseline and the exit surveys, both paper-based interviews, show the highest degree of correspondence (87.1%), followed by responses in the cell phone-based ESM and exit surveys (83.9%), both of which referred to the same time period. Baseline and ESM responses were the least congruent (71.0%) since not only do they differ in mode of data collection they also refer to sexual activity at different time points. Because of small sample sizes, it is difficult to establish congruence patterns with respect to differences in type of exit interview (i.e., whether FTF or SAQ) although, overall, similarity in mode of data collection seems to prevail as a determinant of congruence: baseline FTF interview corresponds better with exit FTF interview, while self-reporting ESM and SAQ surveys also correspond better. Responses on contraceptive use follow these same general patterns.

Table 2. Congruence of responses on sexual activity, by type of survey, gender and martial status

<table>
<thead>
<tr>
<th>Agreement of responses on sex in the last week</th>
<th>Male (n=15)</th>
<th>Female (n=16)</th>
<th>Never married (n=15)</th>
<th>Ever married (n=16)</th>
<th>Total (N=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline vs. ESM</td>
<td>11 (73.3%)</td>
<td>11 (68.8%)</td>
<td>12 (80.0%)</td>
<td>10 (62.5%)</td>
<td>22 (71.0%)</td>
</tr>
<tr>
<td>Baseline vs. Exit FTF</td>
<td>14 (93.3%)</td>
<td>13 (81.2%)</td>
<td>14 (93.3%)</td>
<td>13 (81.2%)</td>
<td>27 (87.1%)</td>
</tr>
<tr>
<td>Baseline vs. Exit SAQ</td>
<td>14 (93.3%)</td>
<td>13 (81.2%)</td>
<td>14 (93.3%)</td>
<td>13 (81.2%)</td>
<td>27 (87.1%)</td>
</tr>
<tr>
<td>ESM vs. Exit FTF</td>
<td>12 (80%)</td>
<td>14 (86.7%)</td>
<td>13 (81.2%)</td>
<td>13 (81.2%)</td>
<td>26 (83.9%)</td>
</tr>
<tr>
<td>ESM vs. Exit SAQ</td>
<td>12 (80%)</td>
<td>14 (86.7%)</td>
<td>13 (81.2%)</td>
<td>13 (81.2%)</td>
<td>26 (83.9%)</td>
</tr>
</tbody>
</table>

Discussion

We used cell phone-based surveys to collect data on sexual activity and contraceptive use among young adults. The software application enhanced real-time collection of data on sensitive topics in the diary method or ESM style. We were able to implement this survey with low direct cost of materials, with mid-market mobile phones and free data collection software.

As with all technology-based data collection tools, this particular mobile platform and methodology did not come without challenges and risks. The learning curve was steep and varied considerably from one respondent to another, despite the fact that ownership and regular use of cell phones was an eligibility criterion for participation. There were occasional reports of data loss (due to difficulty in saving the data) and, in some cases, duplicates to compensate for such loss. We also experienced data loss due to misuse and breaking of one device belonging to a male participant. There were two cases where objections regarding the sensitive nature of the study—specifically the daily query about sexual
thoughts and activities—were raised by people close to the respondents (albeit not the respondents themselves). One case was the American fiancé objecting to his girlfriend’s participation in the study and another was an older brother who protested to the “experimentation” that he said we were subjecting his brother to. In both cases, however, the respondents themselves expressed their wish to continue with the study. Perhaps it is necessary to further emphasize and safeguard the confidentiality of the study. Unfortunately, many cell phone users in the Philippines use their cell phones in some sort of sharing capacity: constraints such as “no load” or “low batt” may require sharing with friends and other household members. In cell phone-based data collection studies, then, such practices must be considered. Despite distribution of devices by the research team, and individual ownership of a cell phone, a phone culture of device sharing may impinge upon confidentiality efforts and perhaps contribute to participants’ reluctance to report on sensitive topics.

Overall, however, participant compliance in the pilot study was quite high, once participants acquired skill in using the technology. The fact that the respondents came from the CLHNS cohort and have established good rapport with OPS and the project staff may have helped guarantee this cooperation and compliance. Nonetheless, a more rigorous training of study participants is recommended. Also, a more thorough and incremental incentivizing scheme needs to be worked out to further enhance participant compliance and commitment, and reduce device mishandling and data loss.

Despite the risks and challenges mentioned above, the pilot study highlights an important finding: that participants show a strong preference for answering on the cell phone rather than through direct interviews. This bolsters the argument for self-reporting and diary methodologies on questions involving sexual activities, especially since participants themselves admit to more honest responses in this modality.

In future studies, we hope to address some of the limitations and concerns cited above. Our next steps focus on issues of scale, developing the pilot scope into a larger study of up to 200 participants in the Metro Cebu area. Additionally, we hope to explore the possibility of surveying couples, i.e., both partners in a married/cohabiting relationship in order to further understand discordance in reported sexual activity.
References:


and the context of sexual coercion in young adults' sexual interactions: Results from a diary study in rural South Africa. Sexually Transmitted Diseases 2006; 33(1):52-58.


