Secular Trends in Response Rates for Controls Selected by Random Digit Dialing in Childhood Cancer Studies: A Report from the Children’s Oncology Group

Greta R. Bunin1,2, Logan G. Spector3,4, Andrew F. Olshan5, Leslie L. Robison6, Michelle Roesler3,4, Seymour Grufferman7, Xiao-ou Shu8, and Julie A. Ross3,4

1 Division of Oncology, Children’s Hospital of Philadelphia, Philadelphia PA.
2 Department of Pediatrics, School of Medicine, University of Pennsylvania, Philadelphia, PA.
3 Division of Epidemiology and Clinical Research, Department of Pediatrics, Medical School, University of Minnesota, Minneapolis, MN.
4 University of Minnesota Cancer Center, Minneapolis, MN.
5 Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, NC.
6 Department of Epidemiology and Cancer Control, St. Jude Children’s Research Hospital, Memphis, TN.
7 Department of Internal Medicine, School of Medicine, University of New Mexico, Albuquerque, NM.
8 Department of Medicine, Center for Health Service Research, Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN.

Received for publication September 19, 2006; accepted for publication January 17, 2007.

Since the mid-1990s, epidemiologists have anecdotally reported difficulty in recruiting controls using random digit dialing (RDD), but few empirical data have been published. From 1982 to 2003, epidemiologists from the Children’s Oncology Group conducted 17 case-control studies using RDD controls. Data for calculating RDD and field response rates were available from eight and 13 of these studies, respectively. Over the period of analysis, the contact rate declined 2.5% per year (95% confidence interval (CI): −3.4, −1.6; p = 0.001), from above 90% in the 1980s to 63–69% in the most recent studies. The response rate (the product of the contact and cooperation rates) showed a decline parallel to that of the contact rate (−2.4% per year, 95% CI: −3.2, −1.6; p < 0.001), from above 80% in the 1980s to 50–67% after the mid-1990s. Field response rates appeared to have declined modestly. The overall response rate (the product of the RDD response and field response rates) paralleled that of the RDD response rate and decreased 2.4% per year (95% CI: −2.7, −2.0; p < 0.001). The current low response rates for RDD indicate a substantial potential for selection bias and a need to seek alternative sources of controls.

case-control studies; data collection; epidemiologic methods

Abbreviations: CI, confidence interval; COG, Children’s Oncology Group; RDD, random digit dialing; SCA, Survey of Consumer Attitudes.

For a generation, epidemiologists have used random digit dialing (RDD) as a source of control subjects. Because the vast majority of US residents have telephones (1), a variety of RDD methods have been used to randomly select controls from the general population. RDD has been commonly used for control selection in epidemiologic research because of its relatively low cost, coupled with acceptable participation rates. Researchers have compensated for its disadvantages, such as the tendency to select persons who spend more time at home, by requiring that randomly selected phone numbers be called at various times of the day and week to maximize the proportion of the population that can be contacted.
Since controls need to have a household phone in order to be selected, having a household phone has become an eligibility requirement for cases as well.

Since the mid-1990s, epidemiologists have anecdotally reported increasing difficulty in recruiting controls (2, 3). For RDD, the perceived decline in response rates is attributed to the widespread practice of screening calls using answering machines, caller identification, and related technology. In addition, an increasing proportion of the population uses cell phones exclusively and does not have “land” lines. Few empirical data have been published that support or refute the anecdotal reports. In an analysis of studies conducted from the 1970s to the 1990s, Olson et al. (4) did not observe a significant decline in the RDD screening response rate or overall response rate among controls selected by RDD. However, they relied mostly on response rates reported by authors, who may have differed in terms of how they calculated the rates. For example, authors may or may not have included persons who could not be contacted in the denominator. Recently, Morton et al. (5) reported that participation rates for population-based controls (not selected exclusively by RDD) declined significantly between 1970 and 2003, decreasing particularly rapidly since 1991.

More recent data obtained from a set of studies with similar methods and standard calculation of response rates could provide evidence to confirm a decline in RDD response rates and to examine whether the contact rate, the cooperation rate, or both have declined. The epidemiologists of the Children’s Oncology Group (COG) have selected controls by RDD in many studies of various childhood cancers, starting in the early 1980s. Case and control parents were then interviewed by telephone. Here we report response rates for RDD screening and telephone interviewing from those studies.

MATERIALS AND METHODS

Epidemiologists associated with the former Children’s Cancer Group and Pediatric Oncology Group and the current COG have conducted 17 case-control studies using controls selected by RDD (6–21). COG is the North American clinical-trials group for pediatric oncology and has as members over 200 institutions that treat children with cancer (22). The Children’s Cancer Group and the Pediatric Oncology Group were two of the four clinical-trials groups that merged to form COG in 2000. For brevity, any or all of the three groups that conducted the studies analyzed here will be referred to as COG.

All of the studies used the same general methods for case ascertainment, RDD, recruitment, and data collection (table 1). Most of the information was derived from published reports, with some details of the methods being obtained through personal communication with the investigators. All cases who received the eligible diagnosis during a specified time interval were ascertained through COG. The criterion for age at diagnosis varied widely, and the upper age limit ranged from 1 year to 20 years. Other eligibility criteria for both cases and controls were the presence of a household phone, availability of the biologic mother for interview (i.e., the child was not adopted or in foster care), residence in the United States or Canada, and physician consent to contact the family. Most studies were restricted to children with English-speaking mothers, but studies E-18 (neuroblastoma), E-21 (medulloblastoma/primitive neuroectodermal tumor), and A0026 (Wilms’ tumor) included Spanish-speaking mothers as well. In two of the more recent studies, namely E-18 (neuroblastoma) and A0026 (Wilms’ tumor), both cases and controls were offered a financial incentive to complete the interview.

For the RDD method used in the majority of the studies analyzed here, part of the case’s phone number served as the base for randomly generating phone numbers. In most studies, randomly generated two-digit numbers were appended to the area code, prefix, and next two digits of the case’s phone number (23). If a control was not found using those 99 phone numbers, RDD continued using randomly generated three-digit numbers appended to the first seven numbers of the case’s phone number. Study E-21 (medulloblastoma/primitive neuroectodermal tumor) used the two-stage Waksberg method within the case’s area code (24). In all of the studies, trained interviewers called each phone number until a respondent was reached or a predetermined number of calls (between 6 and 15) was made, with calls being made on both weekdays and weekends and during both daytime and evening. In addition to matching on part of the phone number, each study required that controls be matched to cases on age or birth date. The age-matching criteria varied among the studies but generally required a closer match in age for younger children. For example, in study E-18 (neuroblastoma), cases aged 3 years or less required a control with a birth date that was within 6 months of the case’s birth date, while older cases required a control birth date that was matched within 1 year. In addition to age, many of the studies matched participants on race/ethnicity, and several matched them on sex. For matching on race/ethnicity, the categories used were either non-Hispanic White and non-White; Black and non-Black; or non-Hispanic White, Black, and other. Most of the studies used individual matching, but the four most recent studies used frequency matching.

Slattery et al. (3) defined contact, cooperation, and response rates to describe participation rates among controls in the RDD and field aspects of studies. For RDD screening, the contact rate is the percentage of households called in which a member of the household answered the phone. (A household phone number is defined as a working number that does not reach a business or other known nonresidence (e.g., a pay phone or nursing home).) The cooperation rate is the percentage of households that participated in the RDD screening among those in which someone answered the phone. The response rate is the percentage of households that participated in the RDD screening out of the number of households called; it is the product of the contact rate and the cooperation rate. The field response rate is the proportion of potential controls who participated in a study interview from among those screened and found eligible.

Contact, cooperation, and response rates were calculated for all studies for which published or unpublished data on outcomes of RDD calls were available. Field response rates were obtained from published reports.

Linear regression was used to examine the change in rates over time with calendar year as the independent variable and rate as the dependent variable. For each study, the last year of data collection minus 1 was used as the study year in the regression. We used the last year minus 1 rather than the median year to account for the fact that the pace of data collection increased over the course of the study. However, using the median year did not change the results appreciably.

RESULTS

Data sufficient to calculate RDD contact, cooperation, and response rates were available for eight studies conducted from 1984 to 2003. Figure 1 presents the changes in rates over that time period. The contact rate declined 2.5 percent per year (95 percent confidence interval (CI): −3.4, −1.6; \( p = 0.001 \)), from above 90 percent in the 1980s to 63–69 percent in the most recent studies. The cooperation rate remained high, at least 86 percent, throughout the time period (change = −0.02 percent; 95 percent CI: −0.7, 0.6; \( p = 0.93 \)). The response rate, which is the product of the contact and cooperation rates, showed a decline parallel to that of the contact rate (−2.4 percent per year, 95 percent CI: −3.2, −1.6; \( p = 0.001 \)). In the 1980s, the response rates were above 80 percent, as compared with 50–67 percent after the mid-1990s. When the analysis was limited to the six studies conducted after 1994, the contact rate showed a steeper decline of 3.1 percent per year (95 percent CI: −5.8, −0.5; \( p = 0.03 \)), the cooperation rate showed an increase of 1.2 percent per year (95 percent CI: 0.6, 1.9; \( p = 0.007 \)), and the response rate showed a decline of 2.1 percent per year (95 percent CI: −4.6, 0.4; \( p = 0.08 \)).

Field response rates were available for 13 of the studies conducted from 1983 to 2003. The field response rates appeared to have declined, although there was some variability in the most recent studies (figure 2). Before 1995, field response rates were 74–80 percent, as compared with 60–76 percent for the more recent studies. The linear decline was 2.0 percent per year (95 percent CI: −2.7, −2.0; \( p < 0.001 \)), the same as that for the RDD response rate. The results did not change when only the eight studies with RDD screening data were analyzed.

DISCUSSION

Our compilation of RDD data from 13 COG studies documents that the contact rate has declined since the late 1980s and is the largest contributor to the decline in the response rate. The cooperation rate remained high during the 20-year period, resulting in response rates that were only slightly lower than contact rates. Participation in phone interviews declined modestly, such that the response rate for the RDD and field portions of the study combined showed a decline which paralleled the decline in the RDD response rate. The studies analyzed were similar in design; all of them ascertained cases through the COG and used similar RDD methods. The studies did differ with respect to matching criteria, age range of the controls sought, details of the RDD method, and investigators. However, the observed declines are unlikely to be attributable to these factors, since none of these characteristics was concentrated in a particular portion of the time period studied. Rather, the observed declines seem to reflect the general RDD method, not some other aspect of the studies.

The decline in the RDD and total response rates observed here is consistent with the findings of Morton et al. (5) for a similar time period (1991–2003). However, the decline they observed, which was based on data from 24 published reports, was about twice that reported here. Morton et al. could not always know how authors calculated the participation rate, since their analysis was based on published reports. In addition, their studies differed more in terms of study population characteristics and control selection methods than those analyzed here, because they analyzed all population-based case-control studies published in one of 10 journals during a defined time period that specified participation rates. It is possible that not all of the population-based studies used RDD controls, although that information was not reported. Although our analysis was based on a small number of studies, the consistency of the control selection methods and study populations suggests that our results are valid for the use of RDD to identify parents for participation in telephone interviews.

Curtin et al. (25) also observed declining RDD response rates in the Survey of Consumer Attitudes (SCA). The SCA has been conducted by phone since 1979, using RDD to identify respondents, and thus provides a resource for observing changes in response rates. SCA contact and response rates dropped between 1985 and 2003, as did the rates in COG studies. The SCA contact rate was approximately 84 percent in 2003, and the response rate was 48 percent. Although the SCA contact rates are higher and the response rates lower than those in COG studies from the same period, a substantial decline was observed in both sets of studies. Unlike the COG experience, the cooperation rate in the SCA decreased from 81 percent to 73 percent. Differences in implementing RDD and in target populations (all persons aged \( \geq 18 \) years for SCA vs. parents for COG studies) may explain the discrepant rates.

Olson (4) analyzed RDD response rates in 13 studies conducted by a single research group at the Fred Hutchinson Cancer Research Center from 1989 to 1996. She observed that RDD screening response rates declined modestly (−0.3 percent/year; \( p = 0.01 \)) but remained greater than 92 percent. In contrast, we observed a much larger decline in RDD response rates over the same time period. Although the COG RDD response rates were similar to those of the Fred Hutchinson Cancer Research Center at the beginning of the 8-year time period studied by Olson, our response rates had declined to approximately 70 percent by the end of that time period. Many factors might explain the difference, including
<table>
<thead>
<tr>
<th>Children’s Oncology Group protocol and cancer type</th>
<th>No. of cases</th>
<th>Years of diagnosis of cases</th>
<th>Years in which study was conducted</th>
<th>Age range (years)</th>
<th>RDD* method (matched portion of phone no., maximum no. of calls made)</th>
<th>Other matching factors</th>
<th>Controls paid?</th>
<th>Spanish speakers included?</th>
<th>RDD response data available?</th>
<th>Field response data available?</th>
<th>Reference no(s.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-01: osteosarcoma</td>
<td>152</td>
<td>1982–1983</td>
<td>1983–1987</td>
<td>0–17</td>
<td>Area code + next five digits, nine calls†</td>
<td>Date of birth (±1 year for cases aged &lt;2 years at diagnosis, ±2 years for older cases)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>6, 23: L. L. R., unpublished data</td>
</tr>
<tr>
<td>E-02: hepatoblastoma</td>
<td>75</td>
<td>1980–1983</td>
<td>Not stated</td>
<td>No age limit (91% of cases were &lt;5)</td>
<td>Area code + next five digits, nine calls†</td>
<td>Date of birth (±6 months for cases aged &lt;2 years at diagnosis, ±1 year for older cases)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>7, 23: L. L. R., unpublished data</td>
</tr>
<tr>
<td>E-03: Ewing's sarcoma</td>
<td>153</td>
<td>1983–1985</td>
<td>1983–1987</td>
<td>0–20</td>
<td>Area code + next five digits, nine calls†</td>
<td>Date of birth (±1 year for cases aged &lt;2 years at diagnosis, ±2 years for older cases)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>6, 23: L. L. R., unpublished data</td>
</tr>
<tr>
<td>E-05: acute nonlymphocytic leukemia</td>
<td>204</td>
<td>1980–1984</td>
<td>1984–1986</td>
<td>0–17</td>
<td>Area code + next five digits, nine calls†</td>
<td>Date of birth (±6 months for cases aged &lt;1 year, ±1 year for cases aged 1–3 years, ±2 years for older cases), race</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>8, 23: L. L. R., unpublished data</td>
</tr>
<tr>
<td>E-07: retinoblastoma</td>
<td>201</td>
<td>1982–1985</td>
<td>1983–1986</td>
<td>No age limit (95% of cases were &lt;5)</td>
<td>Area code + next five digits, nine calls†</td>
<td>Date of birth (±1 year), race</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>9</td>
</tr>
<tr>
<td>E-08: non-Hodgkin’s lymphoma</td>
<td>268</td>
<td>1986–1990</td>
<td>Not stated</td>
<td>0–20</td>
<td>Area code + next five digits, nine calls†</td>
<td>Date of birth (±1 year for cases aged &lt;3 years, ±2 years for older cases), race, sex</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>10, 23: L. L. R., unpublished data</td>
</tr>
<tr>
<td>E-10: rhabdomyosarcoma</td>
<td>322</td>
<td>1982–1988</td>
<td>1984–1989</td>
<td>0–20</td>
<td>Area code + next five digits, nine calls</td>
<td>Date of birth (±1 year for cases aged &lt;5 years, ±3 years for older cases), race, sex</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>12, 13</td>
</tr>
<tr>
<td>E-12: brain tumor (medulloblastoma/ primitive neuroectodermal tumor and astrocytoma)</td>
<td>321</td>
<td>1986–1989</td>
<td>1988–1990</td>
<td>0–5</td>
<td>Area code + next five digits, nine calls†</td>
<td>Date of birth (±1 year), race</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>14; G. R. B., unpublished data</td>
</tr>
<tr>
<td>E-14: acute myelocytic leukemia</td>
<td>525</td>
<td>1989–1993</td>
<td>1990–1994</td>
<td>0–17</td>
<td>Area code + next five digits, nine calls†</td>
<td>Age (±25% of age for cases with a maximum difference of ±2 years), race</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>15, 23: L. L. R., unpublished data</td>
</tr>
<tr>
<td>E-15: acute lymphoblastic leukemia</td>
<td>1,914</td>
<td>1989–1993</td>
<td>1989–1993</td>
<td>0–14</td>
<td>Area code + next five digits, nine calls†</td>
<td>Age (±25% of age for cases with a maximum difference of ±2 years), race</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>16, 17, 23; L. L. R., unpublished data</td>
</tr>
</tbody>
</table>
the skill of the RDD interviewers or the responsiveness of a regional population compared with that of the country as a whole. During the same time period, the Fred Hutchinson Cancer Research Center field response rate for in-person interviews and the COG rate for telephone interviews showed no decline.

The results reported here were based on a small number of studies, which limits the interpretation of some findings. For example, we observed that the field response rate—that is, the willingness of eligible controls to participate—declined only modestly over the study time period. However, the post-2000 data were based on only three studies, which had field response rates of 60 percent (J. A. R., unpublished data), 61 percent (21), and 76 percent (26). The study with the 76 percent rate (26) was influential, and without it the observed decline in field response rates would have been steeper. Without more studies, we cannot know the true change in field response rates. The reported changes in RDD rates were also limited by small numbers. The COG RDD data from the early 1980s were based on only two studies, and thus the true magnitude and/or slope of the decline may differ from that observed. However, regardless of the existence or true magnitude of the decline, the data from the recent studies indicate low contact and response rates. The RDD contact rates in 2002 were approximately 60 percent, the RDD response rates were also approximately 60 percent, and the overall response rates were approximately 40 percent.

The generalizability of our results differs among the various rates studied. The decline in the contact rate is generalizable beyond pediatric cancer studies. People in households that could not be contacted did not know anything about the study, and therefore we expect that those results would apply to studies of any pediatric or adult disease. On the other hand, regional and/or ethnic differences in contact rates may exist, as suggested by the continuing high response rates for studies conducted by the Fred Hutchinson Cancer Research Center, as discussed above. In contrast to the contact rate, we observed that the cooperation rate remained high. Respondents who did not hang up the phone heard information about the study before deciding whether or not to complete the RDD screening. Thus, the disease and age group of interest, as well as interviewer skill and regional/ethnic differences, may have affected the cooperation rate; our results for this outcome may not be generalizable to studies of adult diseases or nonmalignant pediatric diseases.

It has been suggested that participation rates for controls, regardless of source, have declined in recent years (2). These anecdotal reports concern cooperation rates—that is, the willingness to participate once contacted, found eligible, given information, and invited. There are few published data with which to confirm or refute this belief. Morton et al. (5) could not distinguish between the ability to contact controls and the ability to recruit controls once contacted, because that information was often missing from the studies they analyzed. We did not have data on the cooperation rate for the field portion of each study, but we did have the field response rate (i.e., the combination of the contact and cooperation rates). The field response rate declined only modestly, but as discussed above, the most
recent data were based on only three studies, with field response rates between 60 percent and 76 percent, a broad range. Thus, this report and that of Morton et al. (5) do not document convincing declines in participation.

We observed that after 2000, the overall response rate was approximately 40 percent. The lower the response rate the greater the potential for selection bias, since persons who do not answer the phone or complete the interview may differ in characteristics that affect the study results. Some researchers have tried financial incentives to increase the response rate. In 1999, SCA investigators began sending advance letters to all potential respondents for whom a mailing address could be identified (44 percent), since experiments had shown that this practice increased the response rate (25). In 2000, $5 was added to the letter, again based on experiments. These changes did not result in an increased response rate. Thus, although the changes may have prevented a further decline, these approaches are unlikely to increase response rates to previous levels. The small number of COG studies analyzed here, of which three offered financial incentives to participants, did not permit an assessment of the effect of the incentives.

It seems unlikely that the RDD response rate can be increased to previous levels. The inability to characterize nonresponders is the major limitation of RDD and seems insurmountable. Another option might be to consider the ability to be reached by RDD as an eligibility criterion for cases and to collect information from cases on call screening behavior and use of land-line phones. Cases who could not be reached by RDD would be excluded, or the results obtained with and without these cases would be compared. Such an approach may be tenable for common diseases. However, for most pediatric cancers, epidemiologists must contend with small sample sizes, even with nationwide case identification. Thus, if a substantial proportion of cases would have to be excluded because, for example, they did not have land-line phones, those studies would not be feasible.

For pediatric cancer, control groups other than RDD controls must be considered. Birth certificate controls present an attractive alternative (27), as they represent the base population (US births) in which the cases occur. The birth certificate provides information with which to compare responders and nonresponders, and such information can be used to adjust case-control analyses for differences observed between responders and nonresponders. However, the use of birth certificate controls presents challenges as well. In nationwide studies such as those of childhood cancer and other rare diseases, researchers must interact with and receive
approval from many state health departments. In addition, the children identified must be located on the basis of information on the birth certificate. Birth certificate controls have been successfully used in childhood cancer studies in New York and California (28, 29), with approximately 50 percent of parents identified from birth certificates completing the study interview. COG epidemiologists are identifying birth certificate control groups nationwide for an expansion of studies AE24 (infant leukemia) and AEPI04C1 (hepatoblastoma) (30). The feasibility of this control group nationwide and the representativeness of the controls recruited will determine whether birth certificate controls become the standard for pediatric cancer studies.

Declining rates for RDD screening and, to a lesser extent, for participating in a telephone interview are documented here. Persistence of the low response rates reported here and elsewhere will be problematic for epidemiologists. Use of telephone behavior as an eligibility criterion might be useful in reducing the potential for selection bias but implies the possibility of having to exclude a substantial proportion of cases. Based on the data presented here, COG epidemiologists do not plan on using RDD in future studies. Rather, we are exploring alternative sources of controls, especially ones that provide information on nonparticipants with which to assess the potential for selection bias. Birth certificate files may provide a feasible and valid source of controls for studies of pediatric cancers and other pediatric conditions.

ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health grants CA35314, CA36222, CA42479, CA21244, CA29275, CA49450, CA48051, CA57004, CA60951, CA67263, CA75169, CA79940, and CA75385.

The authors thank Greta Anschuetz for her help in producing the graphs. They also thank the hard-working staffs of the studies discussed in this paper.

Conflict of interest: none declared.

REFERENCES


