

REDUCING INTERVIEWER SCREENING AND CONTROLLING SAMPLE SIZE IN A LOCAL-AREA TELEPHONE SURVEY

Martha J. Banks, University of Chicago
Dan E. Hagan, Chilton Research Services

This paper reports on several features of the survey methodology used in two recent telephone surveys -- features that reduced interviewer screening and controlled sample size. An examination of the procedures employed should be useful to many other survey practitioners, illustrating the practical application of methodological principles. The sharing of information about techniques used in complex surveys should help the survey research community to develop and identify optimal procedures for use in future studies.

Approximately five years ago, the Robert Wood Johnson Foundation and the Health Care Financing Administration sponsored a trial program in which five large American cities would establish a network of municipal health centers to provide coordinated primary medical care. To assess the impact of the program, the Health Care Financing Administration, along with the Robert Wood Johnson Foundation, funded two waves of telephone surveys, one conducted in 1980 and the other in 1982. (For brevity and clarity, we will present only the results from the second wave, as both used essentially the same procedures.) Additional information about the study can be found in The Municipal Health Services Program: Improving Access While Controlling Costs? by Gretchen V. Fleming and Ronald M. Andersen, which is the final report to the Health Care Financing Administration and the Robert Wood Johnson Foundation, HCFA 500-78-0097 and RWJF 6798, University of Chicago Center for Health Administration Studies, July 1984.

In each of the five cities, one of the health centers was chosen as the focus of study. A sample of patient addresses, as given on registration forms, was used to determine the service area for use in the survey. Generally, the areas chosen contained about 75 percent of all patients. We estimated that between 5 and 15 percent of all families in each service area contained one or more persons who had been patients of the municipal health center during the preceeding year; while the goal was to have 25 to 30 percent of all family interviews to be conducted with user families. Therefore, quite a bit of screening out of nonuser families would be necessary. In order to reduce the amount of screening the interviewers would have to perform, an oversample of phone numbers likely to belong to user families was employed. The oversample was chosen from a list of phone numbers provided to the facilities by patients during the previous eighteen months. While it was recognized that these lists were not perfect, in that some user families were not on the lists and that many numbers on the lists would belong to nonusers rather than to users, it was expected that oversampling from these lists definitely would be beneficial.

Let us now describe in chronological order the steps used, which will provide more details about the list oversample, as well as present the other

techniques which reduced the amount of interviewer screening and controlled sample size.

SELECTING TELEPHONE NUMBERS

The first step in selecting the sample was to identify the telephone exchanges that existed in each service area. This was done using maps and reverse directories. All streets (and ranges of addresses on these streets) in the service area were identified using several different maps per city. Then reverse directories were used to identify the phone exchanges belonging to persons in these ranges of addresses. Over all five service areas, 88 exchanges were identified for use in the study. The working banks (groups of 100 numbers based on the first two digits after the exchange) were determined next. To the working banks used in the 1980 wave were added any banks listed either in a current reverse directory or in Chilton Research Services' masterfile of exchanges (or both). Once this was done, the universe of eligible phone numbers was complete and sampling could begin.

The set of phone numbers eligible for inclusion in the 1982 sample consisted of all phone numbers assigned in the 1980 random digit sample (regardless of the 1980 outcome), plus a proportionate number of numbers from new exchanges and new working banks, plus the list oversample. The list of phone numbers provided by the municipal health centers was oversampled at 3:1.

This oversampling was accomplished by 1) identifying eligible phone numbers (those on the list in exchanges in the service area) and removing all other numbers, 2) matching the resulting numbers with the regular sample, and 3) from the unmatched list, sampling enough phone numbers so that the sampling rate from the list (including the cases also in the regular sample) was three times the sampling rate of the regular sample. This is illustrated with the diagram shown in Figure 1. The large square (categories 1 through 5) contains all eligible phone numbers; that is, all numbers in working banks in the service area. The smaller square centered within it (categories 2 and 3) represents the regular sample; that is, the sample chosen independently of the list of patient phone numbers. The rectangle containing categories 3 through 6 represents the entire list of patient phone numbers. Category 3 indicates the portion that was chosen independently in the regular sample, category 4 is the additional list sample, 5 is the unselected part of the eligible list, and 6 is the part of the list which is ineligible (because the phone exchanges of these numbers are not in the service area used in the study). The entire sample, then, is represented by categories 2, 3, and 4.

Table 1 provides the number of phone numbers in each of these categories. The number of phone numbers initially sampled ranged from about 13,000 (City 3) to nearly 45,000 (City 5). Of

INTERVIEWING RESPONDENTS
AT THE SELECTED PHONE NUMBERS

Respondents were selected from contacted households by this process: First, street names (and hundred blocks if necessary) were obtained from the respondents so that those living outside the service area could be eliminated. Then, a portion of those households with no member who had used a municipal health center or the municipal hospital during the previous twelve months were screened out, so that only about half the interviews were done with those in nonuser households. The rate at which nonuser households were sampled varied from one third to one half. We started with a rate of two fifths in early replicates; in later replicates the rate was varied as needed.

COMPLETED INTERVIEWS AND RESPONSE RATES

Table 3 shows the distribution of interviewed families by user status and by whether or not the phone number was on the list of patient phone numbers. This table illustrates the importance of the list sample. Had there been no list oversample, the assigned replicates would have yielded the number of interviewed families given by the nonlist portion plus a third of the list sample. For example, if there had been no list oversample in City 1, we would have gotten $(245/3)+61 = 143$ user families and $(185/3)+534 = 596$ nonuser families from the 40 replicates. To obtain interviews with the 306 users actually interviewed in this city's service area, we would have had to have used $306/143=2.14$ times as many random phone numbers. Therefore the use of a list sample was invaluable in obtaining an acceptable mix of cases by user status.

A survey's response rate, which generally is considered to be an important measure of the accuracy of survey estimates, indicates the portion of all eligible cases that actually responded in the survey. In estimating the number of eligible cases, the study's response rate calculations assume that, for categories whose eligibility is unknown, the portion eligible is the same as the portion eligible among those whose eligibility status is known. Each different type of eligibility (household or not, in area or not, selected eligible household or not) was dealt with separately, as was each user-status group. The resulting response rates are presented in Table 4. We hypothesize that the response rates would have been lower if there had been no list sample and no clerical screen because the increased numbers of calls necessary to complete the survey would have reduced the amount of time the interviewers could spend on converting reluctant respondents. Thus, a clerical screen and the use of a list sample reduced the amount of interviewer screening necessary in this study; leading, we believe, to higher response rates and thus to better-quality data. Further, the use of replicates allowed us to control the sample size in an efficient manner.

TABLE 2 RESULTS OF THE CLERICAL SCREEN

| CATEGORY | CITY | | | | |
|--|------|-------|------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 |
| a Additional list sample(not screened) | 949 | 937 | 530 | 670 | 562 |
| b Business or institution | 1274 | 1560 | 764 | 1585 | 1954 |
| c Listed residence; out of area | 774 | 6102 | 820 | 1012 | 3446 |
| d Listed residence; in area | 1229 | 2001 | 1892 | 2243 | 2818 |
| e Unlisted; non-working | 1577 | 2950 | 1575 | 3122 | 3621 |
| f Unlisted; no answer or busy | 1163 | 2176 | 968 | 1851 | 3536 |
| g Unlisted; answer | 1352 | 1436 | 684 | 947 | 2591 |
| a through g = total | 8318 | 17162 | 7233 | 11430 | 18528 |
| a+d+f+g = assigned to interviewers | 4693 | 6550 | 4074 | 5711 | 9507 |

TABLE 3 DISTRIBUTION OF INTERVIEWED FAMILIES BY USER STATUS AND LIST STATUS

| USER AND LIST STATUS | CITY | | | | |
|----------------------|------|-----|------|------|------|
| | 1 | 2 | 3 | 4 | 5 |
| User families | 306 | 276 | 310 | 251 | 327 |
| List | 245 | 222 | 164 | 172 | 112 |
| Not on list | 61 | 51 | 146 | 79 | 215 |
| Nonuser families | 719 | 689 | 713 | 776 | 786 |
| List | 185 | 132 | 111 | 127 | 58 |
| Not on list | 534 | 557 | 602 | 649 | 728 |
| TOTAL | 1025 | 965 | 1023 | 1027 | 1113 |

TABLE 4 RESPONSE RATES BY USER STATUS

| USER STATUS | CITY | | | | |
|------------------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 |
| User Families | 75.2 | 78.9 | 84.8 | 74.1 | 77.0 |
| Nonuser Families | 69.0 | 69.9 | 79.5 | 67.9 | 70.8 |
| TOTAL | 70.7 | 72.3 | 81.2 | 69.2 | 72.6 |