The Survey in Family Practice Research

Kathryn Keller, MSW
Richard N. Podell, MD, MPH
Summit, New Jersey

The survey is a method that can be employed in family practice research. It is an adaptable tool that can be used with good results by those not sophisticated in research. General guidelines for survey research are discussed in this article, and an illustrative example of a survey conducted by the authors is presented.

Recently the Overlook Hospital Family Practice Residency* surveyed its patients concerning satisfaction with their medical care. Having encountered problems in adapting survey methods, we have written this paper in the hope that other family physician-investigators might benefit from our experience. Survey research is an adaptable tool which can be used effectively by persons without sophisticated research backgrounds. However, one must anticipate the standard pitfalls and take reasonable care in adapting the textbook rules to the requirements of the individual research project. Thus, we present general guidelines for the survey as well as an example to illustrate them.

Guidelines for the Survey as a Research Tool

Formulating the Problem

The first and most important step in planning a survey is the clear and specific definition of objectives. Everything else, such as sample, design, and analysis, follows from this. One should ask such questions as, "Does this topic need to be studied?" "Is this topic a manageable size?" "Can we make the problem concrete so that it can be studied?" "What type of research – survey, observation, experimental – would best give us the answer to our problem?" A study to identify bottlenecks in office procedures, for example, will differ greatly from one comparing the rate of compliance with medication directions among patients of two model family practice units.[1]

Selecting a Sample

It is usually impossible and impractical in terms of time, effort, and money to solicit everyone in the population under study. Therefore, a representative sample of the total population is employed. There are several common approaches to defining the survey sample.[2,3] The choice depends largely on the survey goals and the characteristics of the study population. One approach is to assign each member of the target population a number and then select the survey sample from a table of random numbers. When a list of persons in the study population is available (eg, a practice's billing system), it is more convenient to compile the sample by selecting every nth person on the list. The value of n is determined by the required size of the sample, and the first person is selected by a random number falling between one and n. It may be important to be certain that there is adequate representation from critical subgroups of the target population. In these instances, the target population is stratified according to the critical subgroups, and random or systematic sampling is done within each group.

The question of sample size is one which often plagues beginning researchers. Fortunately, for most family practice situations the estimate of sample size can be readily made from formulae found in very elementary texts,[3,4] provided that the investigator can answer three questions:

1. What is your estimate of the actual variability of the target population?
2. How close must the results from the survey sample be to the frequencies which actually occur in the target population?
3. How confident must you be in the answer to question #2? (Usually being right 95 times in 100 is good enough.)

Choosing the Instrument

The researcher must decide which survey instrument will best serve the study's objective – the self-administered questionnaire or the interview. Each has advantages and disadvantages. The self-administered questionnaire is generally less expensive since it does not require the training or paying of interviewers. It is easy to standardize and distribute, and the results can be more easily compared. It can be completed anonymously, an advantage which is sometimes crucial.

On the other hand, the questionnaire's relative inflexibility is an important disadvantage. There is no way for the researcher to assess whether the respondent understands the questions. If a mailed questionnaire is used, response rates will be

*Overlook Hospital is a 541-bed suburban community hospital which serves a generally middle-income population. The Family Practice Residency cares for approximately 4,000 patients in two model family practice units which are run as a private, fee-for-service, group practice.

From the Division of Family Practice, Overlook Hospital-Columbia College of Physicians and Surgeons, Summit, New Jersey. Requests for reprints should be addressed to Ms. Kathryn Keller, Research Associate, Overlook Family Practice Associates, 193 Morris Avenue, Summit, NJ 07901.

lower, 30 to 50 percent is typical, and therefore the study design must incorporate techniques to test whether those actually answering the questionnaire are representative of the entire population sample surveyed. Obviously the respondent's literary (and visual) acuity must be assumed!

The interview provides a better response rate since the interviewer can make repeated attempts to contact the respondent until an interview or a refusal is obtained. There is a potential for more feedback from the respondent and a greater certainty that each question is understood. The interviewer can probe more deeply into areas which seem promising.

Unfortunately, the interview is also time-consuming, and if paid interviewers are used, costly. Volunteers can be effective, but require initial time investment for training. Some researchers[5] feel it is unwise to rely on volunteers because their interest may lag, they may lack certain specialized interviewing skills, and high standard of work performance cannot be used as a disciplinary measure. Interviews cannot be anonymous, of course, and there is always the possibility that the interviewer may consciously or unconsciously affect the respondent's answers. Different interviewers may not be comparable in the way they question or report. Unless the interview is highly structured, it may be difficult to compare the respondents' opinions or to translate them into quantitative terms.

It is possible to use a combination of techniques at different stages of a study. For example, an interview schedule can be used in the earlier stages to help develop a relevant standard questionnaire. Or non-respondents to a self-administered questionnaire can be interviewed in person or by phone. Fortunately, it is generally possible to combine answers from personal and self-administered questionnaires.[6]

Constructing the Questionnaire

Whether a self-administered questionnaire or interview is chosen, the researcher must be concerned about detailed form and content.[1-3]

Each question must meet certain criteria. The question should be clear to the respondents and mean the same to them as to the researcher. Frequently, respondents confuse or misinterpret questions because terms and concepts which seem clear to the researchers are not so to the respondent. This is especially important advice for physicians, whose attitudes and knowledge on health topics almost always differ from their patients, no matter what the patient's background. Clues to poorly written questions include low response rates to the question and responses differing markedly from what was expected.

Each question should refer to one idea so that it will be clear to what the person is responding. Questions should not be leading, ie, suggesting that one answer is better than another. Finally, questions should be relevant and related to the research objectives.

The type of question should be appropriate. There are two types of questions: the closed (or fixed alternative) question and the open question. A closed question is designed so that it contains a number of mutually exclusive and exhaustive answers among which the respondent is asked to choose. The answers to closed questions are easier to code and quantify for analysis. However, closed questions can be artificial and force choices inappropriately, which may confuse or irritate respondents.

An open question is unstructured, and the respondent can answer without the constraint of predetermined choices. Thus, it allows for unanticipated responses which may reveal significant information. However, it is more difficult to code. In interviewing it is more time-consuming, and more interviewer skill is required. It may also be more difficult for the interviewer to record responses accurately.

Open and closed questions can be combined. A very broad open question can be followed by specific closed questions in the same area. Or specific questions can be followed by open questions which deal with the intensity of attitudes or the level of information.

If the interview is used, its structure must be considered. There are three main formats: the standardized, the semi-standardized, and the non-standardized interview. The standardized interview asks the same questions in the same order and manner to all respondents. The standardized interview can be used in situations where the subjects of the survey are homogeneous and the survey covers matters which are factual or non-emotive.

The semi-standardized interview is less rigid. Instead of a predetermined schedule of questions, the semi-standardized interview is based on a list of required questions which can be asked in any sequence, and the questions are formulated in language appropriate for each respondent. The semi-standardized interview is useful in situations requiring the collection of data from groups which are heterogeneous and which cover matters which are emotional or sensitive.

In the non-standardized interview there is no predetermined questionnaire and no checklist of required information. A close medical analogy is the psychiatric interview. The content of the interview may vary from one respondent to another. In its most formal style, the focused interview, it approaches the semi-standardized interview. In the focused interview, the interviewer is required to guide the interview in such a way that all the predetermined topics are covered in the interview, but there are no set ways to go about it. The objective of the focused interview is to center attention on a particular set of topics about which the interviewer wishes to know the respondent's views.

Different interview formats can be employed at different stages of a study. A focused interview might be used in the early stages of a study when the interviewer is exploring the possibility of a variety of research topics. The standardized or semi-standardized interview can then be used after the researcher has learned more about his study population.

The organization of the questionnaire is important. It must be organized so that it can easily be handled by the respondent, the interviewer, and the investigators. From the respondent's point of view, the sequence of questions should be logical and *non-threatening.* From the interviewer's point of view, the questionnaire should be designed for maximum ease of administration. It should be easy to read the questions and record answers. From the researcher's point of view, the questionnaire should be designed to facilitate editing, coding, and identification of the questionnaire.

The questionnaire should be pilot-tested before administering it to the

study population.[1,3] Impatience at this critical step is common, but failure to adequately pilot-test is an invitation to disaster. Those pretesting the questionnaire should be representative of the ultimate target population. They should be encouraged to suggest changes and to voice their reactions to specific questions and to the questionnaire as a whole. One should go over these questionnaires carefully to spot trouble signs, for example, items left blank or yielding useless information or misinterpretations. Appropriate additions, deletions, and modifications can then be made.

Dealing with Non-Response

If the self-administered questionnaire is to be used, one must deal with the problem of non-response.[2,3,7] The initial response to a self-administered questionnaire depends upon such factors as the characteristics of the population to whom the questionnaire will be sent, the degree of interest in the subject of the survey that can be aroused, and the prestige of the sponsor of the survey. A cover letter should explain the study and the importance of the respondents' contribution. It should set a reasonable but firm return date. The letter should be written on official stationery and signed by someone who can lend prestige. The questionnaire should be pleasing to the eye and appear as though it will not take too much time to complete. Its instructions must be easily understood on one reading. There should be a stamped, self-addressed envelope.

However, even with excellent planning, rarely will more than one half the sample respond to the initial solicitation. Thus, a second or even a third contact will be required for non-respondents. Follow-up may be by letter or telephone. It should be non-punitive and assume the tone that the respondent had intended to return the questionnaire but perhaps overlooked doing so. If your first solicitation responses were signed or coded for individual identification, the follow-up is easier. If responses are unsigned and if anonymity has been guaranteed, then follow-up will require recontacting some who actually responded to the first solicitation. Occasionally, those who have already responded are disturbed by the second contact, but usually they are not.

The critical question, of course, is "Are the non-respondents likely to be different from the respondents in their answers?" And if so, is the group of non-respondents large enough to affect your conclusions? In a mailed questionnaire, even with a second and third solicitation, the non-responders often remain a substantial part of the total sample. Occasionally their opinions will differ from those of responders. Indeed, in some instances their opinions may well be closely linked to the reasons why they chose not to respond.

If the issues are important enough, one can follow-up on each non-respondent and interview him by telephone or in person. This is very time-consuming and usually not practical. However, there are other techniques which, while less rigorous, are much more efficient. One can simply compare the known characteristics of the respondents (eg, age, race, utilization patterns) with those of the entire sample. If, for example, one racial group is severely under-represented among the respondents, probably there is a serious bias being introduced. If all groups have responded proportionately, one is a little more confident of the representativeness of the results. A better technique is to compare the actual responses of each wave of follow-up to see if the respondents who are harder to reach give opinions which differ from those in the first wave. Most researchers rigorously take a small but randomly selected number of non-respondents and make a major effort to interview one hundred percent of this sample. If this is done successively, the subsample of non-respondents will represent the opinions of all non-respondents for most practical purposes.

Processing the Data

After the data is collected, it must be processed.[3] Questionnaires are edited for completeness, clarity, accuracy, and uniformity. If there is missing or confusing information, it may be possible to estimate the answer from other information available on the questionnaire. Or the interviewer will be able to fill in a gap from memory (a good reason to edit the same day). If not, the respondent should be contacted to clarify a confusing or missing response.

Next, responses and questions must be grouped into meaningful categories. This process is called coding and its purpose is to facilitate analysis. Often, coding is automatic as with, for example, the variable of *sex*. Occasionally, questionnaire choices listed separately are combined because they are, for the purposes of the study, similar in meaning (excellent, very good, good) or because they are so infrequent that they can be grouped as "other." Coding is especially important (and difficult) in less structured interviews where highly individualistic responses must be grouped by their common major themes if any statistical analysis is to be made.

Many questions can be pre-coded on the interview or self-administered questionnaire. A designated code number can be printed next to each possible response. This facilitates the tabulation of the frequencies of responses. If punch cards are used, coding on the questionnaires also facilitates the translation of the data to the card.

The difficulty involved for the coder in classifying the responses into categories will depend on the soundness of the code categories. If the categories are not well constructed, the coder will continually find answers which either do not fit anywhere or which might fit several categories.

Tabulation is the next step after the questionnaires are edited and coded. Tabulation amounts to nothing more than a counting of the number of responses falling into the different categories for each question. Tabulation can be done either manually or by machine. The decision to use machine or hand tabulation depends on several factors. If the sample is small ($n<300$), then it is easier to tabulate by hand since machine tabulation requires punching and verifying cards. If data analysis is straightforward (frequencies or simple cross-tabulations) hand tabulation is fine. However, if multiple variable correlations are needed, hand tabulation becomes too slow and there is a greater chance for error. Mechanical card sorting may increase efficiency for analysis of intermediate complexity. For major projects, electronic computers may be used. Machine tabulation is more costly, perhaps too costly for most family practices, but it has one great advan-

tage – its flexibility. When one is not certain what tabulations are needed, machine tabulation is preferable because an unlimited number of analyses can be performed.

Analysis

After the data has been processed, the results are analyzed, which can be done at several levels.[3] Analysis need not be statistical. In *problem identification,* even a small well-documented incident can be very important. (This level is somewhat analogous to the medical case report.) A more involved analysis is *quantitative,* eg, describing the frequency of responses, the mean, average, range, and dispersion. Most sophisticated is *statistical* inference where information from the sample is used to estimate the characteristics of the entire target population and to make inferences about relationships between variables. Much practical statistical inference can be done with the aid of an elementary statistics textbook. More elaborate (eg, multivariate) analysis requires statistical consultation, which should be done *before,* not after the study is designed, lest the study's validity be destroyed by failure to attend to a statistical fine point along the way. These studies will be the exception. Usually you can design your own. However, it never hurts to have someone comfortable with statistics or epidemiology review your protocol before you get too far.

Use of Results

Finally, research, if it is to have any meaning, must be interpreted and used, whether the application is global or for local purposes only. As an example, research findings can be effectively used administratively. They can be translated into judgments concerning program success or failure and lead to modifications of the program. The survey can be repeated at intervals to measure and note change.

An Illustrative Example

The authors' survey can serve as an illustration of the use of these guidelines. Many of the problems that developed and the decisions we rendered are of the type other family practitioners might encounter.

In our residency program, we have used various evaluative methods to determine if we are meeting our standards of quality medical care. Chart audits and reviews, and resident and staff evaluations are standard. But until the survey was conducted, we had never looked at our program from the point of view of the patient. Although the staff had heard complaints or compliments from time to time, feedback was intermittent. We had no idea what the typical patient felt about us. A systematic, in-depth approach was needed to study patient satisfaction with the program.

Since "satisfaction" is rather nebulous, we needed to make the concept operational. Overall satisfaction would be determined by ratings that patients gave our staff and office procedures. We hypothesized that the extent and degree of deficiencies felt by the patient would affect our delivery of service and create dissatisfaction. In turn, dissatisfaction with the program could prevent growth of our practice and damage its acceptability to patients and to the community. We were aware of some deficiencies but did not know their extent; others would be identified by the patients. In addition, patient comparisons of our practice to other providers would give us an idea of where we needed to improve.

Thus, our survey had three purposes: documenting several already known problems (eg, parking, in order to persuade the hospital of our need for assistance), identifying previously unknown problems with personnel and office procedures, and learning how our patients felt about our strengths and weaknesses in comparison to other providers of medical care.

How we chose our sample was influenced by our goal of problem identification. Therefore, we had to be certain our sample would represent opinions from both model family practice units as well as from patients who were most likely to be dissatisfied. We stratified the target population and chose equal numbers from each model unit. From each office, half the sample was chosen from among people who were our regular patients and half from among those who had not become our regular patients. The latter group was selected out of proportion to their contribution to our visit-volume, on the assumption that they would view our efforts more critically. Within each group, systematic sampling was used.

In developing our instrument, we used both the interview and the self-administered questionnaire. First, a focused telephone interview was conducted. We would introduce ourselves and our purpose and let the patient take it from there. After the patient had spontaneously vented his feelings, we moved the conversation into areas we thought might be pertinent. These interviews made it easier to write a standard questionnaire by sensitizing us to the patients' concerns (eg, in the beginning we asked about inpatient care, forgetting that few of our patients have been hospitalized), the ways they interpreted our language, and their own typical styles of expression. The self-administered questionnaire was chosen as the main instrument because of its convenience and minimal expense.

Before mailing the written questionnaire, we pilot-tested it by interviewing a small group of patients who were not included in the study sample. Those pilot-testing the questionnaire made suggestions, and the questionnaire was revised. Even with pilot-testing, certain flaws in the questionnaire became evident only after responses started coming in. For example, we asked patients to rate their satisfaction with parking facilities on a four-point scale. However, a simpler question, "Have you ever had any problem parking?" probably would have related better to our research objective (ie, proving there is a problem).

Despite our efforts to ensure responses, our return rate was 40 percent. The data collected from the 40 percent implied that generally we were providing an excellent service. Unfortunately, with such a low return rate, we could not be sure this favorable impression was a valid one. We did not know if the 40 percent accurately represented the entire patient population. It was possible that nonrespondents had a different and unfavorable perception of the practice.

In order to follow-up on nonrespondents, we used a telephone, semi-standardized interview. All questions from the self-administered questionnaire were asked, but the order and wording were left to the discretion of the interviewer. Follow-up could have been conducted by another mailing of questionnaires, but we felt

in this case the interview should be used because of its effectiveness in eliciting a response. A trained volunteer was to interview 100 percent of a small sample of non-respondents (non-respondents were easy to identify since almost all the respondents had signed their questionnaires). However, as it turned out, our volunteer interviewer was very competent and enthusiastic. Therefore, she eventually interviewed all non-respondents to the questionnaire, not just a sample. In retrospect, it might have been as easy to interview all patients selected for survey. However, we suspect that our luck with a volunteer interviewer was probably exceptional in this case.

Processing and tabulation of the data was relatively simple. Analysis also remained elementary. Since our major goal was problem identification, significance of results was not always determined statistically but sometimes in administrative terms (this is similar to the important exceptions one might find in a chart audit). On the other hand, the difference between respondents and non-respondents to the first solicitation, while at times significant statistically (ie, not likely to occur by chance), was not different enough to be significant in terms of the administrative decisions we had to make.

Our results were to be utilized for administrative purposes, and we learned much about our problems. For example, we had not suspected that there was a problem with our billing system prior to the survey. As a result of the survey we were able to improve the billing procedure. In the case of parking, we knew there was a problem and used the survey findings to document the problem for the hospital administration. Most important, we found that our model family practice unit was highly regarded compared to alternative sources of care. This gave us confidence in the decision to expand the residency program with the (correct) assumption that the practice would continue to grow.

Discussion

Because of its flexibility, use of the survey for research is feasible in most family practices. It can be used for a variety of purposes. Depending on the demands of the practice in terms of time, money, and effort, one can choose a self-administered questionnaire, a telephone interview, or an in-depth personal interview. Analysis can be simple or sophisticated. The survey can be adapted to fit the capabilities of the researcher as well.

We have dealt selectively with the survey as a research tool. More detailed information can be found in a number of reasonably straightforward introductory texts. What remains is for the family physician to plan the objective of his research.

References

1. Sellitz C, Johoda M, Deutsch M, et al: Research Methods in Social Relations, revised ed. New York, Holt, Rinehart and Winston, 1959
2. Burton TL, Cherry G: Social Research Techniques for Planners. London, George Allen and Unwin Ltd, 1970
3. Moser CA: Survey Methods in Social Investigation. London, William Heinemann Ltd, 1959
4. Parten M: Surveys, Polls and Samples: Practical Procedures. New York, Cooper Square Publishers Inc, 1966
5. Parten M: Op cit, p 135
6. Sudman S: Reducing the Costs of Surveys. Chicago, Aldine Publishing Co, 1967, pp 46-57
7. Good CV, Scales DE: Methods of Research. New York, Appleton-Century-Crofts, 1954, pp 626-628