SURVEYS OF PHYSICIANS: AN OVERVIEW

Sara L. Thran, American Medical Association, Marc L. Berk, Project HOPE
Sara L. Thran, American Medical Association, 515 N. State Street, Chicago, IL 60610

KEY WORDS: physician surveys, methodological research, review

This paper provides a critical analysis of the significant findings from methodological research conducted on surveys of physicians and discusses potential areas for future research. Lessons learned from physician surveys should prove useful to researchers interested in studying other types of medical providers, as well as the larger audience of researchers interested in surveying small establishments with professional personnel whose time is valuable.

Physician surveys differ from general population surveys, and are often particularly difficult because of the nature of physicians and their practices. In surveying physicians, one must be aware that physicians have demanding work schedules and that their time is valuable. Thus, there is a high opportunity cost for physicians' time. Additionally, physicians are frequently approached for surveys and thus are often reluctant to participate. They typically have receptionists or other "gatekeepers" who are vigilant in protecting them from calls by interviewers.

As Parsons et al. (1993) note, telephone interviews of physicians demand special consideration because the respondents usually can only be reached by first securing access through individuals, often referred to as gatekeepers, who are responsible for screening the physicians' calls. The interviewer must convince the gatekeeper that it will somehow be to the advantage of the physician to participate or at least get the gatekeeper to persuade the physician to listen to the request.

Traditionally, surveys of physicians have been viewed as one type of survey of elite professionals (Sudman, 1985). With recent changes in the medical practice environment, however, it has become increasingly apparent that physician surveys are a type of <u>business establishment survey</u>. Frequently the physician practice represents a business in which a staff of physicians share joint resources. Many physicians are employees of the practice who may not be authorized to reveal information about the practice. As more physicians enter managed care settings, interviewers must understand that they often are not contacting a solo practitioner who is free to make independent decisions about his/her use of time.

There has long been consensus that physician surveys are inherently difficult (see for example Martin, 1974). The reason for this sentiment is understandable. Overall response rates on physician surveys tend to be lower than those on household surveys and physician surveys frequently are expensive. It should be recognized that it is not always interviewing the physician, per se, that makes the survey difficult. Often physician surveys focus on information that is extremely detailed and difficult to recall. In addition, many physician surveys frankly are boring, perhaps not to researchers but certainly to the physician. To the health policymaker, it is very important to understand the types of practice costs incurred by the physician. Most physicians, however, do not find it interesting or enjoyable to be asked a detailed series of questions about how much specific items of medical equipment have depreciated over the last year. Providing detailed, factual information on practice characteristics is difficult and time-consuming and, for many physicians, intrusive. When confronted with such a survey, many will refuse to respond and others agree only after several months of persuasion.

It is important to keep these differences in complexity in mind because the general perception about the difficulty of physician surveys may unnecessarily discourage certain types of inquiry. In fact, a large number of surveys of physicians have been conducted at modest cost and have obtained high response rates. Well-designed surveys that focus on physician attitudes about the practice of medicine can obtain high response rates without any phone followup and without even considering more elaborate procedures, such as payment of monetary incentives. Short questionnaires certainly are more likely to be completed than longer ones. While experienced survey firms often have considerable expertise in obtaining high response rates, frequently it is the nature of the survey itself rather than the organization conducting it that is responsible for obtaining high rates of compliance. Survey firms that conduct large numbers of physician surveys often find considerable variance in the response rates obtained on different types of surveys administered to physicians, as seen in Exhibit 1.

Because obtaining physician cooperation is especially difficult for predominantly economic surveys, research on techniques for increasing response rates to such surveys has recently been conducted. Much has been learned about surveying physicians through several large-scale ongoing or periodic surveys such as the

American Medical Association's Socioeconomic Monitoring System and government-sponsored surveys including the Physician's Practice Cost and Income Survey, and the National Medical Expenditure Survey. This body of research has not been widely disseminated but will be reviewed here. This paper provides a synthesis of significant findings from recent methodological research on physician surveys, particularly the large economic surveys. Areas of research to be reviewed include:

- Importance of high response rates;
- Reasons for survey nonresponse:
- Methods to reduce response burden/reduce nonresponse; and
- Adjustments for nonresponse.

It is becoming more important than ever before to develop accurate methodologies for obtaining data from physicians. As key decision makers in the health care system, physicians are of considerable research interest. The current debate over health care reform offers exciting new opportunities for health survey researchers. During the last few years we have seen a proliferation of new health policy initiatives proposed by the Administration, Congress, as well as by the private sector. The behavioral response of physicians to health care reform is critical in understanding the strengths and weaknesses of each proposed initiative. Although survey methodologists have a long tradition of studying the ability of patients to provide accurate data about their health care experiences, less attention has been focused on the accuracy and validity of physician surveys.

DESCRIPTIONS OF MAJOR PHYSICIAN SURVEYS

Socioeconomic Monitoring System

The American Medical Association's Socioeconomic Monitoring System is a series of annual telephone surveys of non-federal patient care physicians (excluding resident physicians). Data are collected on medical practice characteristics such as hours worked, number of patient visits, fees, Medicare participation, practice income and expenses. Each survey collects data from approximately 4,000 physicians through an interview averaging 25 minutes in length. The survey is conducted between March and July. Each survey includes reinterviews with respondents who were initially interviewed a year earlier, as well as interviews with physicians selected for the first time. Survey response rates have ranged between 65 and 70 percent. The surveys have been conducted by an outside survey firm (Mathematica Policy Research or the RAND Corporation).

Physicians' Practice Cost and Income Survey (PPCIS)

The Health Care Financing Administration has sponsored several national surveys of physicians since the mid-1970's. These telephone surveys have collected data pertaining to physicians' practice patterns, productivity, practice costs, income, and Medicare participation. The most recent of these, the 1988 PPCIS, obtained responses from approximately 3,500 non-federal, patient-care physicians (excluding residents), who were either owners of their practice or were employed by another physician or group of physicians. The survey response rate was 61%; the average interview length was 45 minutes. The 1988 PPCIS was conducted by the National Opinion Research Center, between July 1989 and March 1990.

National Medical Care Expenditure Survey - Physician Practice Cost Survey

In 1977 the National Center for Health Services Research (now called the Agency for Health Care Policy and Research) conducted the Physician Practice Cost Survey as a component of the National Medical Care Expenditure Survey. The survey essentially was a shortened version of the Physicians' Practice Cost and Income Survey conducted by HCFA. This, however, was one of the few federally sponsored surveys where the sample was not generated from the AMA Masterfile. The PPS was a follow-up survey to the National Medical Expenditure Survey where household respondents were asked to provide the names of their physicians. It was hoped that the PPS could then be linked to the data obtained on the household survey to develop accurate estimates about what types of physicians served what types of patients. Although the overall response rate was 74%, only about half of all respondents could be linked to completed questionnaires from **all** of the physicians identified during the NMES survey.

IMPORTANCE OF HIGH RESPONSE RATES

Although there is a clear consensus among survey researchers that high response rates are critical, the studies that have been conducted on non-response patterns in physician surveys provide some evidence that physician surveys are somewhat more resilient to the effects of non-response bias than other types of surveys. Studies have been conducted comparing survey respondents and non-respondents. It is also possible to examine difficult/late vs. easy/early respondents to determine whether the additional efforts to increase response rates are warranted.

Berk (1985) examined results from the 1977 Physician Practice Cost Survey to see whether single variable estimates would have been different if the survey period had lasted 2 months or 4 months instead of the actual 6 months. The findings suggested that the addition of late responders to the sample did not affect most estimates of key demographic variables. When Berk compared the estimates after 4 months of the field period, when the response rate was at 49%, to the final survey estimates, after a 74% response rate was reached, only 3 of the 14 estimates for key provider and practice characteristics changed more than 5% and only 1 estimate (the numbers of doctors in practice) changed by more than 10%. Seven of the estimates had changed less than 1% as the response rate moved from 49% to 74%.

With the Socioeconomic Monitoring System (SMS) survey, Thran, et al. (1986) found that means and regression coefficients are not affected by the addition of late respondents and other difficult-to-interview sample members. More recent studies of other physician surveys by Guadagnoli and Cunningham (1989) and Sobal and Ferentz (1989) similarly found that the addition of late respondents did not lead to a more representative sample.

Rosenbach and Ammering (1992) compared early and late respondents to the 1988 Physicians' Practice Cost and Income Survey. They defined late respondents as those who responded within the last month of the field period. They found that the probability of late response varied systematically in several respects. Busier physicians had a higher probability of late response. There were differences across practice arrangements, with the probability of late response increasing with practice size, as well as for physician employees. Higher earning physicians were more likely to be late respondents.

Descriptive studies which have compared the demographic characteristics of non-responding and responding physicians generally have found they are similar for most important demographic characteristics (Goodman and Jensen, 1981; Berk and Myers, 1980; Kaspar, 1979; Loft, 1980).

Marder and Thran (1989) examined the impact of survey nonresponse in the SMS survey. A representative subsample of nonrespondents to the 1988 SMS survey was recontacted several months later and asked to complete an abbreviated telephone survey to collect information on key survey variables, raising the response rate for the subsample from 69% to 86%. The results from respondents to the special survey were generally quite similar to those of the regular respondents. Thus, Marder and Thran concluded that additional survey efforts designed to increase the response rate did not appear to be warranted.

These studies should not lead us to conclude that appropriate efforts to maximize response rates are unnecessary. They do suggest, however, that using the unit response rate as a "rule of thumb" general measure of survey quality is overly simplistic. Survey researchers need to be careful in determining when to end the field period because of the systematic differences between early and late respondents. In particular, it may take extra time for the busy, high earning physicians, especially those in large group practices, to obtain permission or to take the time to participate. It is important to compare cumulative estimates obtained early in the field period to cumulative estimates obtained later. Even if there are significant differences in the responses obtained at the end of the field period, unless the number of late responses is large enough to significantly change the overall cumulative estimates, continuing the field period is often not advisable.

REASONS FOR NONRESPONSE

Almost 20 years ago, Martin (1974) noted that, "as members of a high prestige profession, physicians' opinions are constantly being solicited, and as a result, physicians are plagued with surveys." Today physicians are working in larger practices, often under the direction of others. They may be increasingly sensitive to external pressures to use their time efficiently and cooperating with survey researchers may be considered infeasible.

Sudman (1985) noted four reasons for nonresponse to surveys of professionals. They are: professionals are busy and their time could be better spent on other professional tasks; the value of the survey is not clear; confidentiality concerns; and the individual questions appear biased or do not allow for the full range of answers. All of these are still applicable as reasons for nonresponse to surveys of physicians.

Some researchers are concerned that the role of the federal government in regulating health care may also eventually lead to lower rates of survey participation. There may be some physicians who believe that refusing to participate in a federal survey is a vehicle for registering displeasure with federal policies.

Studies of reasons for non-response have been conducted by Berk and Myers (1980) on the 1977 Physician Practice Survey and by Sprachman (1985) on the 1983 Physician Practice Cost and Income Survey. Berk et al. (1989) compared reasons for nonresponse on the 1988 National Survey of Physicians with the reasons noted in the earlier two efforts. All three were federally sponsored surveys on physician costs that were conducted by very large survey organizations. These studies compared the reasons for nonresponse noted on the non-interview response forms completed by the interviewers. On all of these surveys, the non-response forms were developed as field mechanisms to determine the best way of converting refusals. They were not intended to provide comprehensive or uniform data about reasons for nonresponse.

Concern with the federal government did not appear to be a major factor in decisions on survey participation. Hostility toward government did not appear to increase as a reason for nonresponse between the 1977 and 1983 surveys. This finding may be misleading since the most hostile physicians may be those who refused to even discuss the survey with the interviewer. The number of cases in which the interviewer was not even able to ascertain a reason for refusal doubled from 1977 to 1983 but declined dramatically between 1983 and 1988. On the 1988 National Survey of Physicians, almost half of all refusing physicians specifically noted that they did not have time or were too busy to participate.

In the SMS survey, nearly one-third of the refusals come from receptionists rather than directly from the physicians. A large proportion of survey nonrespondents are not outright refusals but are "noncontacts." For example, some physicians request that a mail questionnaire be sent rather than completing the interview over the telephone, and they never return the mail survey. Other physicians do not keep appointments for their interviews.

METHODS FOR REDUCING RESPONSE BURDEN/REDUCING NONRESPONSE

A number of methods have been used to increase response rates to surveys of physicians and to reduce response burden. These include: use of physician callers; allowing the use of "proxy" respondents designated by the physician; use of monetary or nonmonetary incentives; use of endorsement letters; and using a combination of data collection methods. The findings on each of these methods are summarized below.

Use of Physician Callers

Bostick, et al. (1992) describe a new approach to obtaining a high response rate in a physician telephone survey. As part of the Minnesota Heart Health Program, a telephone survey of physicians was conducted regarding changes in physician practice for the prevention of cardiovascular disease. Physicians who participated in the 1987 survey were resurveyed in 1989. The 1987 and 1989 surveys contained identical sets of 32 questions relating to management of risk factors for cardiovascular disease.

For initial recruitment in 1987, physicians were mailed an introductory letter. Their offices were then telephoned by survey staff to schedule an appointment for a 20-minute telephone interview. Those refusing or whose office staff denied access to the physician were sent a second letter urging reconsideration and were then retelephoned by a study physician whose purpose was to schedule an interview. All interviews were conducted by survey staff - none were done by the study physicians. Identical procedures were used in 1989 for those who completed an interview in 1987.

Follow-up calls by an investigating physician on the 1989 survey increased the response rate from 66% to 92%, resulting in the addition of 69 respondents. This involved approximately 40 hours of physician caller time.

At this point, it is difficult to assess how well this procedure would work in a large national study of physicians. Obviously, it would be quite costly to use physician recruiters in a large study, but it may be a valuable technique to use late in the field period or for particularly reluctant respondents.

Use of Proxy Respondents

In telephone surveys, it is common for a member of the sample to designate another individual to complete some or all of the interview. The individual who completes the interview for the sample member is called a proxy respondent.

The decision whether to respond directly to a survey or to use a proxy has an economic component. Anyone contacted about participating in a survey will weigh the advantages and disadvantages of participation. A part of this calculation, whether explicit or implicit, is to determine the least cost method of responding. The real cost of responding to a survey depends on the amount of time involved for the respondent and the value of that time. Designating a proxy is one way for a physician to minimize the cost of participating by substituting lower cost personnel time for more expensive physician time. Alternatively, the physician might recognize that some of the survey questions could be answered more easily by a business manager, receptionist, or accountant who routinely handles some of the business aspects of the practice. In that case, it is not the value of the time but the differential cost of information retrieval that matters.

Many physician surveys are designed to be administered directly to the physician, with the use of proxies strongly discouraged. However, often when the field period is drawing to a close and an adequate response rate has not been reached, rules about proxy respondents are relaxed. The quality of the responses on physician surveys may be enhanced if it is explicitly recognized during the design of the survey that, under certain circumstances, responses from proxies are acceptable. Researchers should design the survey so that, if necessary, proxies can serve as suitable respondents.

Item nonresponse and perhaps survey nonresponse can be minimized by providing physicians the option to designate a proxy to respond to one or more sections of the survey. However, the value of proxy respondents depends on their ability to respond on behalf of the physician. Several studies have been conducted on the usefulness of proxy reporting in economic surveys of physicians.

Berk, et al. (1981) examined the usefulness of proxy reporting in the 1977 PPS. Approximately 34% of the cases utilized proxy respondents. Their findings suggest, that while it may be preferable to interview the physician directly, proxies can provide estimates for most of the important variables in the survey. In general, the interviewers felt that greater cooperation was obtained from physicians than from proxies. Proxies had very low item response rates to physician net income and also had some difficulty answering other "personal" questions on the physician's age and medical training. However, proxies and physicians had similar item response rates to questions on hours worked, patient visits, fees, and other practice characteristics. Generally, data collection from proxies was found to be of lower cost and higher efficiency than data collection from physicians.

Of the respondents to the 1987 SMS survey, approximately 15% used proxies for some or all of the interview. Marder and Thran (1988) found that those who designated proxies had higher predicted hourly wages than those who did not designate proxies and that hourly wage was a significant predictor of use of a proxy. SMS respondents who designated proxies differ significantly in term of their demographic characteristics from those who responded directly to the survey. Proxy respondents had lower item response rates to physician net income and expense questions but had higher response rates for the fee section of the survey.

Schneider, et al. (1992) analyzed proxy effects in the 1988 PPCIS. In this survey, about 43% of the physicians designated a proxy for the cost section. Schneider, et al. examined characteristics of physicians who used proxies, item response rates for proxy and non-proxy cases, and differences in data reported by proxy and non-proxy cases. They found personal, geographic, and practice differences among those who did vs. did not designate a proxy; in particular, physicians in larger practices, those who accepted Medicare/Medicaid patients, and who were busier were more likely to use proxies. Hourly wage was a significant predictor of use of a proxy in this survey. Proxy respondents had similar item response rates to those of physician respondents. Proxy respondents were found to report significantly higher values on most expense items; even after controlling for practice characteristics, cases using proxies had higher practice costs. However, it was not possible to determine which group provided the more accurate information.

To summarize, the use of proxies offers a number of benefits in surveys of physicians. The use of proxies helps to minimize item nonresponse and perhaps even survey nonresponse. It is reasonable to assume that some of the physicians who designate proxies might not respond to the surveys if they did not have this opportunity to shorten their own time to complete the interview. The use of proxies may minimize guessing on the part of physicians on certain practice-level questions.

Use of Incentives

Promised incentive payments have been shown to improve response rates to physician surveys. It is not well understood why payments are effective in improving response rates. Gunn and Rhodes (1981) speculate that physicians view incentive payments as remuneration for their time, but physicians place a value on their time that far exceeds the amount of even the most generous payments for interviews.

Advance payments have also been demonstrated to improve response rates in physician surveys. Advance payments may be an effective way to signal the importance of the research, by using a rather bold and unusual gesture. According to Berry and Kanouse (1987), social exchange theory also suggests that advance payments should be effective, in that advance payment initiates an exchange transaction that many individuals are likely to feel obligated to complete. The effectiveness of advance payment depends on the perceived legitimacy of the request, and on the subtlety with which it is presented.

Research available to date has not substantiated the effectiveness of nonmonetary incentives to improve response rates to physician surveys.

Gunn and Rhodes (1981) experimented with paying physicians for participating in a 15 to 30 minute telephone interview on attitudes toward influenza immunizations. They tested promised payments of $0, $25, and $50 with general/family practitioners, internists, pediatricians, and industrial physicians. Only 40 physicians were assigned to each payment level, but the size of the incentive payment affected response rate - the response rates were 58%, 69%, and 77% for the $0, $25, and $50 levels. Sensitivity to payment level varied by specialty, with general and family physicians the most sensitive to payment level. Weber, et al. (1982) found that these same levels of incentives affected physicians' response rates to a personal interview.

Berry and Kanouse (1987) compared the effectiveness of a $20 advance payment with a $20 promised payment in a lengthy (20-32 pages) mail survey of 2,147 physicians. The questionnaire was designed to measure physicians' familiarity with the National Institutes of Health Consensus Development Program and assess the extent to which physicians followed procedures recommended by NIH consensus panels in their own practices. The advance incentive group had a response rate of 78% and a refusal rate of 11%; the promised incentive group had a 66% response rate and a 14% refusal rate. Some specialties were more sensitive to the timing of payment than others. Prepayment had little effect on overall characteristics of respondents obtained. Prepayment proved to be efficient, in that it boosted early response rates, reducing the number of cases requiring follow-up.

A variation of the Berry and Kanouse experiment was conducted by Berk, et al. (1993). This study examined whether or not the decision to provide a monetary incentive could be postponed until the results of an initial mailing were available. This would enable researchers to restrict the use of monetary incentives to surveys where early results indicate response rates may be inadequate. The findings confirmed the results of Berry and Kanouse and indicated that the use of a prepaid incentive leads to a large increase in response rates. The study indicated, however, that the value of a prepaid incentive is greatly diminished if the incentive is postponed until a second mailing. This study also found that almost 90% of non-respondents do not cash the prepaid incentive checks.

Since the prepaid incentive reduces the needs for additional mailings and callbacks, the cost of the incentive is greatly mitigated. Thus the prepaid incentive represents an effective and relatively inexpensive way to reduce survey non-response.

Mullen et al. (1987) tested the impact on the response rate to a national mail survey of family physicians of including a small nonmonetary incentive (sticker with slogan) in the first mailing. They found no effect on either the initial or final response rates.

In 1992, a small experiment was conducted in the SMS program on the use of nonmonetary incentives. A random subset of the sampled physicians were sent a card offering the choice of an emergency flashlight or shower radio to be sent to the physician or a staff member that the physician designated. Unpublished analyses indicated that the incentives were not effective, and their use was discontinued.

To summarize, the use of incentives is still relatively new in physician surveys. Monetary incentives have not been used in the AMA's SMS survey or in the large government surveys. The hesitation to use incentive payments in these surveys is, of course, partly due to cost considerations; a payment of even $20 to every respondent would greatly increase the cost of the survey. In addition, at least with AMA's survey, there is concern about how physicians would perceive an incentive payment to participate in the survey (and particularly how member physicians would feel about AMA dues being used for such a purpose). However, given the encouraging results obtained in smaller scale surveys, either advance or promised incentive payments may be more seriously considered for AMA-sponsored and government-sponsored surveys, if response rates decline dramatically in the future.

Endorsement Letters

Most of the major physician surveys utilize endorsement letters from professional associations. Although the use of endorsement letters is accepted practice, there has been little work to examine the effect of the letters on response rates in physician surveys.

Loft (1981) discussed the methodology of the National Ambulatory Medical Care Survey. Post-survey evaluation interviews showed that over half the respondents were influenced by the endorsement letters.

The 1993 SMS survey is using, for the first time, specialty society endorsement letters in combination with the AMA letter. Although no experimental design was used, some inferences may be made on the effect of the letters on response rates and/or number of calls needed to complete the interview.

Combination of Data Collection Methods

In order to maximize response rates, most physician surveys use a combination of data collection methods. At a minimum, generally telephone and mail are used. Certified mail and in-person interviews have been used to a lesser extent.

Ogborne and Rush (1986) reported on an experimental study comparing two methods for dealing with initial nonrespondents to a mail survey of health and social service professionals: a second mailing or a telephone interview. Individuals who had not responded four weeks after the initial mailing were assigned to one of the two treatment conditions. For those in the telephone group, most questionnaires were actually received in the mail. The cost per additional completed questionnaire was approximately $13.00 for the telephone group and $6.66 for the second mailing group. Thus, they concluded that the second mailing was more cost effective than attempts to conduct telephone interviews.

Mullen et al. (1987) examined three types of follow-up contact to a national mail survey of family physicians. Two weeks after the initial mailing, a reminder postcard was sent; two weeks later a letter and questionnaire were sent; and two weeks later the same materials were sent by certified mail. They found that the response rate was increased substantially by the second and third follow-up contacts (11% and 15%, respectively). The non-labor costs for each returned questionnaire were approximately $2.00 for the post-card reminder, $7.34 for the second mailing, and $6.25 for certified mail.

Olson, Thran, and Strouse (1987) examined the effect on overall response rate and compare characteristics of respondents to special data collection efforts in the SMS program, including the provision of a mail questionnaire and numerous callbacks. Different types of physicians were found to respond to the special efforts, and those who respond to special efforts are different from the cases obtained without special effort. Repeated callbacks were found to be very effective in increasing the survey response rate. Mail questionnaires were the most costly of the special efforts examined and only a small number of completed cases were added by provision of mail questionnaires.

To summarize, a combination of data collection techniques is generally necessary in order to obtain acceptable response rates in physician surveys. Mail surveys often involve telephone follow-up, just as most telephone surveys eventually involve the use of a mail questionnaire if particular respondents can only be obtained through that mode. It would be advisable for survey designers to recognize that some respondents are likely to have a preference for a mode of interview other than the primary one. If other interview modes are incorporated into the original survey design, differences in responses between mail and telephone are likely to be minimized.

Summary

It should be acknowledged that not all physician surveys will respond similarly to efforts to raise the response rate. The use of financial incentives should probably be restricted to those surveys that are unusually burdensome to the physician. Surveys that are interesting to the physician and inflict a relatively low response burden can often obtain adequate response rates with more modest efforts. Many projects simply lack the resources to consider the use of a significant monetary incentive and researchers should be aware of the more modest methods that have been reported to be helpful. Mullen et al. (1987) obtained high response rates from additional mail contacts as did Ogborne et al. (1986) and Shosteck and Fairweather (1979). Even procedures as simple as using a stamped, rather than a franked, return envelope or a business reply envelope, have resulted in significant increases in response rates (Shiono and Klebanoff, 1991; Urban, et al., 1993). Parsons, et al. (1993) found that experienced interviewers were more successful than inexperienced; interviewers in conducting physician surveys, in terms of bypassing physicians' gatekeepers.

NONRESPONSE ADJUSTMENTS

Unit or survey nonresponse

Survey nonresponse can be viewed as a significant problem if the number of nonrespondents is large or if those who do not respond to the survey are systematically different from survey respondents. For surveys that achieve relatively high response rates, standard weighting strategies are generally acceptable ways to adjust for unit nonresponse when there is some information about the characteristics of the nonrespondents.

For example, in adjusting SMS results for survey nonresponse, weights are derived by first dividing the AMA Physician Masterfile population of non-federal, patient care physicians and the survey respondents into 200 cells defined by specialty (10 categories), years since graduation (5 categories), AMA membership status (2 categories), and board certification status (2 categories). Unit response weights are constructed as the ratio of the number of physicians in the population to the number of SMS respondents in each cell.

All weighting strategies depend ultimately on the assumption that nonrespondents are similar to respondents when adjustments have been made for their known characteristics. Marder and Thran (1988) report the results of an experiment conducted to test the validity of this assumption within the SMS survey program. By surveying a random sample of survey nonrespondents and examining their characteristics and responses to key questions, Marder and Thran found evidence to support the continued use of standard weighting strategies.

Item nonresponse

Within otherwise completed questionnaires, certain items may have missing responses due to refusal or insufficient knowledge on the respondent's part. If one believes that the answers for the data items would be distributed in a similar manner for respondents and nonrespondents, there is no need to be concerned about the occurrence of missing data. In this situation, the available data could be analyzed directly and reliable estimates obtained for most survey statistics. However, the distribution of respondent data is often very different from the nonrespondent distribution. Therefore, survey estimates obtained from respondent data will be biased with respect to describing characteristics of the survey population unless compensations are made for the missing data. For this reason, imputation procedures are often implemented to reduce the bias caused by missing data.

For item nonrespondents, information often exists that can be used to predict the missing response. Several methods commonly used to deal with item nonresponse are:

- mean value imputation, which replaces missing data with the average of the respondent data;
- hot deck imputation, which uses data from similar respondents to replace missing values;
- regression (or model-based) imputation, which uses respondent data to model the question response in terms of other survey data and then replaces missing data with the model prediction.

Thran and Gillis (1992) applied each imputation method listed above to two key SMS variables with particularly low response rates, annual practice expenses and annual net income. The "adjusted" results were compared with "nonadjusted" results. A simulation was performed to assess the possible impacts that each of the imputation methods may have on key summary statistics. The hot-deck and regression approaches generated reasonable sample statistics. However, when complex relationships between variables were examined (as in a log earnings regression model), only the model-based approach worked fairly well, and the no-imputation approach was the best.

The use of imputation techniques in physician surveys is not universally accepted at present. The 1988 PPCIS uses regression imputation. Imputation is not used on the SMS data, although it is being considered, at least for public use versions of the data file.

AREAS FOR FURTHER RESEARCH

Given the changing medical practice environment and physicians' increasing reluctance or inability to participate in lengthy, complicated economic surveys, survey researchers must be more flexible in designing physician surveys. Some of the techniques described above such as use of physician callers, incentive payments and allowing the use of proxy respondents will probably be increasingly necessary in order to obtain acceptable response rates. In addition, it will probably become critical for survey length to be minimized and only "interesting/relevant" questions to be included on physician surveys (with proxies used for other parts of the survey). Continued methodological research will be needed and the results should be widely disseminated in order to benefit all health policy survey researchers.

A major problem in the evaluation of physician surveys is the absence of a well established source of data for verification. There are dozens of validation studies that have been conducted with household respondents on health surveys. Frequently, after asking the respondents questions about their illness patterns, use of services, and expenditures for care, we conduct followup surveys or administrative recordchecks to see whether respondents actually knew the type of conditions for which they were treated, or whether they could accurately report their use and expenditures for services. Over the past 30 years, there have been numerous investigations examining the National Health Interview Survey and the accuracy of household reports; over a 15 year period, there have been investigations and evaluations of the accuracy of data reported as part of the 1977 and 1987 National Medical Expenditure Survey. Generally, in such investigations the administrative record or followup study is viewed as the standard to which the household reported data is compared. This enables us to conduct studies on what types of respondents are able to report accurately and what types of survey procedures lead to better response.

The health survey methodologist interested in physician surveys, however, has no comparable standard with which to compare answers. This has severely inhibited our ability to talk about the accuracy of information that has been collected. Therefore, most of the methodological work discussed here and elsewhere focuses on issues such as response rates and methods of improving field efficiency. Much less is known about whether or not the physician actually has provided an accurate answer to the question. New opportunities, however, are emerging which may allow us to examine how well the physicians are able to report at least certain types of phenomena.

Recently there has been a great deal of interest in physician payment reform. New payment scales attempt to tie reimbursement rates to, among other factors, the amount of time and effort a physician devotes to specific medical procedures. Current reimbursement rates have been influenced by the answers physicians have given on surveys asking them about the amount of time they spend on specific procedures. In the future we expect to see a significant number of studies comparing the results obtained from physician surveys to those obtained from physician logs, or patient records in which the time spent on inpatient procedures is noted. The area of physician practice costs, and practice patterns, is more difficult. Conducting validity studies about the information obtained on physician practice cost surveys would appear to be problematic in the absence of some type of audit procedure unlikely to be accepted by a physician respondent. Questions on issues such as the number of patients on Medicare or Medicaid who are being served by particular physicians, however, may be subject to verification as public payers develop better administrative record systems for monitoring which patients are going to which types of providers. Thus, there is reason to believe that these advances may make conducting validity checks more feasible.

CONCLUSION

A recurring theme throughout this conference is the importance of reducing the burden of responding to establishment surveys. We believe this is particularly important in the rapidly evolving environment of the physician.

Health survey researchers must recognize the need for new approaches in the design of physician surveys and make an appropriate investment in meeting these new challenges. The need for new information about the behavioral response of physicians to new policy initiatives is likely to increase. Therefore, we must develop technologies and procedures to enable us to obtain accurate data in this more complex environment.

NOTE: The authors gratefully acknowledge the comments made by Sandra Berry, Director of the Survey Research Group at RAND.

REFERENCES

Berk, M. Interviewing Physicians: The effect of improved response rate. American Journal of Public Health, Vol 75, No. 11, 1985.

Berk, M. and S. Meyers. Reasons for nonresponse on the Physicians' Practice Survey. Proceedings of the Social Statistics Section, American Statistical Association, 1980.

Berry, S. H. and D.E. Kanouse. Physician response to a mailed survey: An experiment in timing of payment. Public Opinion Quarterly, 51:102-116, 1987.

Bostick, R. et al. Using physician caller follow-ups to improve the response rate to a physician telephone survey. Evaluation and the Health Professions, 15:420-433,1992.

Goodman, L. and L. Jensen. Economic survey of medical practice: AMA's Periodic Survey of Physicians, 1966-1978. In: Proceedings of the Third Biennial Conference of Health Survey Research Methods. Rockville, MD: National Center for Health Services Research, 1981.

Guadagnoli, E. and S. Cunningham. The effects of nonresponse and late response on a survey of physician attitudes. Evaluation and the Health Professions, 12:318-328, 1989.

Gunn, W. and I. Rhodes. Physician response rates to a telephone survey: Effects of monetary incentives on response level. Public Opinion Quarterly, 45:109-115, 1981.

Harkins, E. An evaluation of the reliability of data gathered from three primary care medical specialties using a self-administered log diary. In: Proceedings of the Third Biennial Conference of Health Services Research. Rockville, MD: National Center for Health Services Research, 1981.

Kasper, J. Problem respondents. In: Andersen, R., et al., Total Survey Error. San Francisco, CA: Josey Bass, 1979.

Loft, J. Methodology of the National Ambulatory Medical Survey. In: Proceedings of the Third Biennial Conference on Health Survey Research Methods, Rockville, MD: National Center for Health Services Research, 1981.

Marder, W. and S. Thran. Proxy or principal: Cost minimizing strategies of survey respondents. Proceedings of the Survey Research Section, American Statistical Association, 1988.

Marder, W. and S. Thran. Are nonrespondents different? Results of an experimental design. Proceedings of the Survey Research Section, American Statistical Association, 1989.

Martin, B. Don't survey physicians: American Medical Association, 1974.

Mullen, P., et al. The cost-effectiveness of randomized incentive and follow-up contacts in a national mail survey of family physicians. Evaluation and the Health Professions, 10:232-245, 1987.

Ogborne, A., et al. Dealing with nonrespondents in a mail survey of professionals. Evaluation and the Health Professions, 9:121-128, 1986.

Olson, L., Thran, S. and R. Strouse. The effectiveness and costs of special data collection efforts in a telephone survey of physicians. Proceedings of the Survey Research Section, American Statistical Association, 1987.

Parsons, J., et al. The effect of interviewer characteristics on gatekeeper resistance in surveys of elite populations. Evaluation Review 17:131-143, 1993.

Rosenbach, M. and C. Ammering. Comparison of Early and Late Responders to the 1988 Physicians' Practice Cost and Income Survey. September 1992.

Schneider, J., et al. Analysis of proxy effects in the 1988 Physicians' Practice Cost and Income Survey. August, 1992.

Sheatsley, P.B. and J. Loft. Comment on Gunn and Rhodes. Public Opinion Quarterly, 45:571-572, 1981.

Shiono, P. and Klebanoff, M. The effect of two mailing strategies on the response to a survey of physicians. American Journal of Epidemiology, 14:539-543, 1991.

Shosteck, H. and W.R. Fairweather. Physician response rates to mail and personal interview surveys. Public Opinion Quarterly, 43:206-216, 1979.

Sobal, J. and K. Ferentz. Comparing physician responses to the first and second mailing of a questionnaire, Evaluation and the Health Professions, 12:329-339, 1989.

Sudman, Seymour. Mail surveys of reluctant professionals. Evaluation Review 9:349-360, 1985.

Thran, S. and K. Gillis. A comparison of imputation techniques in a physician survey. Proceedings of the Survey Research Section, American Statistical Association, 1992.

Thran, S., W. Marder and R. Willke. Probability of response: A multivariate probit analysis. Proceedings of the Survey Research Section, American Statistical Association, 1986.

Urban, N., et al. Effects on response rates and costs of stamps vs. business reply in a mail survey of physicians. Journal of Clinical Epidemiology, 46:455-459, 1993.

Weber, S., et al. The impact of two clinical trials on physician knowledge and practice. Market Facts, Inc., 1982.

Exhibit 1

Response Rates to Several Physician Surveys

Government Sponsored Surveys (NORC)

1977	Physician Practice Survey (NCHSR)	74.0
1983	Physicians' Practice Cost and Income Survey (HCFA)	69.0
1986	Patient Identified Physician Survey Pretest	89.0
1988	Physicians' Practice Cost and Income Survey	61.0

Private Surveys (Gallup)

1988	Knowledge, Attitudes and Behaviors of Physicians Concerning Health Care Services	70.0
1987	Mission and Organization Climate Study	75.0
1991	Mission and Organization Climate Study	75.0
1991	Attitude Study of Physicians Regarding Organizational Climate	65.0
1992	Physician Study Regarding Methodologies to Diagnose and Treat Chlamydia	80.0

ESTABLISHMENT SURVEYS OF THE NATIONAL CENTER FOR HEALTH STATISTICS

Thomas McLemore and W. Edward Bacon, National Center for Health Statistics
Thomas McLemore, NCHS, 6525 Belcrest Road, Room 952, Hyattsville, Maryland 20782

KEY WORDS: Health surveys, data collection

Introduction

The National Center for Health Statistics (NCHS) is the United States Federal government's principal agency for vital and health statistics. As one of the Centers in the Centers for Disease Control and Prevention in the Department of Health and Human Services, the NCHS collects data covering the full spectrum of health-related events from birth to death. Currently, the NCHS has more than a dozen data systems that collect information through vital records, personal interviews, physical examinations and laboratory tests. The NCHS also conducts a number of health care establishment and provider surveys which collect information on the utilization of health care resources in the U.S. This paper provides an overview of the background, development and current status of the NCHS family of establishment surveys, collectively called the National Health Care Survey (NHCS).

NCHS collects data directly from the health care establishments or facilities, or from their records, rather than the recipients of care, because they are cost-effective sources for identifying health care events, such as hospitalizations, surgeries and long-term stays, which are relatively rare events in the general population and because they can provide the most accurate and detailed data on diagnosis and treatment,and on the characteristics of the institution. Data are obtained through a variety of mechanisms, including abstraction of medical records, completion of patient encounter forms, compilation of data from states and professional associations, and purchase of data from commercial abstracting services. These general purpose data are used by policy-makers, planners, researchers and others in the health community to monitor changes in the use of health care resources, to monitor specific diseases and to examine the impact of new medical technologies.

Background

Through the early 1980's, NCHS provided data concerning the use of health care resources primarily through four data collection efforts: three national probability sample surveys, the National Hospital Discharge Survey (NHDS), the National Ambulatory Medical Care Survey (NAMCS) and the National Nursing Home Survey (NNHS); and an inventory of health facilities, the National Master Facility Inventory (NMFI). These studies were designed to operate continuously or on short periodicity cycles and to provide data on the health care settings where the bulk of medical care was provided in the 1960's and 1970's, i.e., in the inpatient wards of hospitals, in physicians' offices and in nursing homes. However, during the past decade, dramatic changes have taken place in the organization, financing and delivery of health care in the U.S. Among the more important factors affecting the health care delivery system are:

> Cost containment--In response to rising health care expenditures, which increased from $250 billion in 1980 to $666 billion in 1990, both public and private purchasers of care have moved toward the institution of payment reform. The prime example of Federal reform is the Medicare Prospective Payment System. The debates on health care reform indicate the critical nature of this issue.
>
> Medical effectiveness--Legislative and departmental initiatives have focused on the issues of medical effectiveness and outcomes of care in attempts to improve the efficiency and effectiveness of clinical practice.
>
> Aging population--With the growing number of older persons in the U.S. population and an increasing life expectancy has come an increased demand for health and social services to support this elderly population.
>
> Medicine and technology--The development and dissemination of new diagnostic and treatment modalities has shifted a number of procedures from the inpatient to outpatient and ambulatory settings. This movement has promulgated the emergence and growth of a variety of new facilities to address this health care market.

The factors indicated above lead to a greater diversification in the delivery of health care as seen in the proliferation of insurance and benefit programs; the increased concern regarding the adequacy and cost of long-term care services; changes in the organization of health care institutions; new physician practice arrangements; a shift from inpatient to outpatient care, from office-based care to other ambulatory settings, and from nursing homes to a host of long-term care settings; and growth in the type and number of alternative sites of care.

The National Health Care Survey

In 1988 in an attempt to better address the data needs of this changing health care environment, the NCHS began to restructure its establishment and provider-based surveys in a program which is called the National Health Care Survey (NHCS). The objectives of the NHCS are to:

- provide nationally representative data on the use of health care resources for the major sectors of the U.S. health care delivery system.
- produce data on an annual basis, thus eliminating data gaps and addressing the seasonality and epidemiology of disease.
- base the surveys on an integrated cluster sample design. This design includes a first stage sample of Primary Sampling Units (PSUs) which is a geographic area usually defined by a county or group of counties. This design provides for operational convenience and efficiency. The PSUs used in the NHCS are the PSUs, or a subset of the PSUs, used in the National Health Interview Survey (NHIS), one of the NCHS population-based surveys.
- develop the capability of conducting patient follow-up studies to examine issues related to the outcome and subsequent use of medical care.

Currently, the NHCS is composed of eight survey activities. These include, in the area of hospital and surgical care, the National Hospital Discharge Survey and the National Survey of Ambulatory Surgery; in the area of ambulatory care, the National Ambulatory Medical Care Survey and the National Hospital Ambulatory Medical Care Survey; in the area of long-term care, the National Nursing Home Survey and the National Home and Hospice Care Survey. Also included are the National Health Provider Inventory and the NHCS Patient Follow-Up Studies.

In the area of hospital and surgical care, the NHCS includes the National Hospital Discharge Survey (NHDS) and will soon include a survey of visits to freestanding and hospital-based ambulatory surgery centers.

The NHDS is a national probability survey of discharges from non-Federal, short-stay hospitals which was initiated in 1965 and conducted continuously since that time. The 1965 NHDS sample was selected based on a two-stage design with stratification by bedsize and geographic region using the NMFI as the sampling frame. Within each sample hospital, a systematic random sample of discharges was selected. Until 1985, NHDS data were collected by a manual system of discharge sample selection and data transcription. In 1985, data tapes containing discharge medical abstracts were purchased from commercial abstracting services and discharge samples were selected from those tapes.

In 1988, the NHDS was redesigned utilizing a three-stage sample design based on a subsample of the NHIS PSUs as the first stage of selection and emphasizing the collection of medical record data in machine-readable form. The target population for the NHDS includes discharges from noninstitutional hospitals, exclusive of Federal, military and Veterans Administration hospitals. Hospitals whose specialty is general (medical or surgical), maternity or children's general or short-stay hospitals (i.e., hospitals with an average length of stay for all patients of less than 30 days) are included in the survey. In addition, hospitals must have six or more beds staffed for inpatient use.

Within a first stage sample of 112 PSUs, a sample of approximately 540 hospitals was selected with probability proportional to size (as measured by the annual number of discharges) after stratifying by hospital specialty/bedsize class and abstract service status. The SMG Hospital Data File was used as the sampling frame. The annual hospital response rate is approximately 92 percent.

Annually, a systematic sample of approximately 250,000 discharges is selected. Discharge data are collected using one of two methods: manual abstraction of data from medical records by hospital staff or by field staff of the Bureau of Census, or the purchase of machine-readable data from state data systems, commercial abstracting services, or directly from hospitals. Approximately 70 percent of the sample discharges are obtained in an automated form.

The NHDS discharge data set conforms to the Uniform Hospital Discharge Data Set (UHDDS) and includes the medical record number; admission and discharge dates (from which length of stay is calculated); date of birth; sex; race; ethnicity; marital status; expected sources of payment; discharge status; disposition; diagnoses; surgical and diagnostic procedures; and dates of procedures. Hospital characteristics, such as current ownership and bedsize, are also collected.

To complement the inpatient data from the NHDS, the NCHS is developing the National Survey of Ambulatory Surgery, a survey of freestanding and hospital-based ambulatory surgery centers. A feasibility study completed in 1991 investigated technical and methodological issues, such as the development sampling frames, a data set and data collection procedures. A pretest of data collection instruments and procedures for the national survey was conducted in the Spring of 1993. The National Survey of Ambulatory Surgery will be initiated in 1994.

In the area of ambulatory care, the NHCS includes the National Ambulatory Medical Care Survey and the National Hospital Ambulatory Medical Care Survey.

The NAMCS is a national probability survey of visits to non-Federal physicians who are principally engaged in office-based patient care practice, but not in the specialties of anesthesiology, pathology or radiology. Telephone contacts and nonoffice visits are excluded. NAMCS was conducted annually from 1973 through 1981 using a three-stage survey design with samples of PSUs, physicians and office visits. The NAMCS was conducted again in 1985 using a three-stage sample design with more detailed physician specialty stratification at the second stage of selection and optimized numbers of physicians selected from these specialty strata.

In 1989, the NAMCS was reinstituted on an annual basis using a subset of the NHIS PSUs as the first stage of selection. The second stage consists of a stratified sample of eligible physicians selected from the master files maintained by the American Medical Association and the American Osteopathic Association. Each year, a sample of approximately 2,500-3,000 physicians is drawn from the universe of about 300,000 physicians. About 75-80 percent of the sample physicians agree to participate.

The third stage is the selection of visits within the annual practices of sample physicians. The physician sample is divided into 52 random subsamples of approximately equal size; then each subsample is randomly assigned 1 of the 52 weeks in the survey year. A systematic random sample of visits is selected from the physician's practice using a patient log or register. Physicians or physicians' staffs complete patient encounter forms for approximately 30 visits over the one-week reporting period. About 45,000 visits are sampled annually.

The 1992 patient encounter form includes the date of visit; date of birth; sex; race; ethnicity; expected source(s) of payment; referral status; whether the visit was injury-related; whether the patient smoked cigarettes; principal complaint(s), symptom(s) or other reason(s) for visit; diagnoses; whether the patient was seen before and whether the condition was new; a disease checklist; ambulatory surgical procedure(s); diagnostic/screening procedure(s); therapeutic services; medications; disposition; and duration. Physician characteristics, such as specialty and type of practice, are collected in an induction interview.

To complement data from the NAMCS, the National Hospital Ambulatory Medical Care Survey (NHAMCS) was fielded in 1992 to provide data on the utilization of ambulatory medical care services in hospital emergency and outpatient departments. Development of the NHAMCS has been underway for more than a decade. Research conducted in 1977 and 1984 provided information on the availability of data items, appropriate data collection procedures and construction of sampling frames. Work completed in 1990 included the development of the national sample design and a pilot test of data collection forms and procedures.

The NHAMCS is based on a four-stage sample design. Within 112 PSUs, a nationally representative sample of 600 non-Federal, short-stay hospitals (using the same hospital definition as the NHDS) is selected based on probability proportional to the size of the emergency and outpatient department (as measured by the annual volume of visits) after stratifying by type of service. Sample hospitals are randomly divided into 16 panels. Each panel is assigned a four-week reporting period; therefore, each sample hospital rotates into the NHAMCS once every 15 months.

Within each sample hospital, the emergency department and up to five outpatient clinics are selected. The outpatient clinics are selected with probability proportional to size based on the expected number of visits. A systematic sample of visits is selected from patient logs or registers. Data are collected by physicians or hospital staff, or in about 15 percent of the hospitals by field staff from the Bureau of the Census, who complete a patient encounter form for each sampled visit. Hospitals are given the option of collecting data prospectively, i.e., at or near the time of the visit, or retrospectively from the medical record. Annually, about 75,000 visits are sampled.

The patient record forms used for the emergency and outpatient departments are similar to the NAMCS form and include the date of visit; date of birth; sex; race; ethnicity; expected source(s) of payment; principal complaint(s), symptom(s), or other reason(s) for visit; diagnoses; diagnostic and screening services; medications; disposition; and providers seen. Additionally, the emergency department form contains major reason for visit; cause of injury; urgency; whether the visit was alcohol- or drug-related; and procedures. And the outpatient department form asks whether the patient was referred; whether the patient was seen before and whether the condition was new; ambulatory surgical procedure(s); and therapeutic procedures.

In the area of long-term care, the NHCS includes the National Nursing Home Survey (NNHS) and the National Home and Hospice Care Survey (NHHCS).

The NNHS was conducted in 1973-74, 1977 and 1985 to obtain information on nursing homes, their expenditures, residents and staff. Planning is underway to conduct the next NNHS in 1995.

The 1985 NNHS collected information from a stratified probability sample of 1,220 nursing homes of the 20,000 nursing homes in the coterminous U.S. Facility information was collected through interviews with the nursing home administrators and expense information was collected through self-administered questionnaires completed by the facility's accountant or administrator. The facility and expense questionnaire response rates were 88 and 68 percent, respectively.

Facility data include size, ownership, Medicare and Medicaid certification status, staff, services and financial characteristics. Data was collected on a sample of approximately 5,400 current residents and 6,300 discharged residents. The current and discharged resident questionnaires included resident's date of birth; sex; race; ethnicity; marital status; place of residence prior to admission; health status; functional status; services received; monthly charges; source of payment; and, for discharges, the outcomes of care. The information was collected through a personal interview with the nurse responsible for the resident's care who referred to the medical record to obtain the requested the information. Data about registered nurses was collected through a self-administered questionnaire from a sample of about 3,500 registered nurses. The response rates were 97, 95 and 80 percent for the current resident, discharged resident and registered nurse staff questionnaires, respectively.

The NNHS also included a telephone survey of the resident's next-of-kin, usually a family member, who was asked to provide information about the resident's financial status, functional status prior to admission, reasons for admission and prior nursing home use.

In 1992, the National Home and Hospice Care Survey (NHHCS) was initiated as an annual survey to collect data on home health agencies and hospices, and their current and discharged patients. This survey was based on a feasibility study conducted in 1990 and a pretest of data collection forms and procedures conducted in 1991.

The NHHCS utilizes a three-stage stratified sample design based on a first-stage sample of 198 PSUs and a second-stage sample of approximately 1,500 hospices and home health agencies selected from the 1991 National Health Provider Inventory. A sample of approximately 9,000 current patients and 9,000 discharge patients is selected from these agencies.

Agency data are collected through personal interviews with agency administrators. Facility data include ownership; Medicare/Medicaid certification; number and type of patients; services provided; and number and type of employees. Patient data are obtained from personal interviews with caregivers who consult the patient's medical record. Patient data include date of birth; sex; race; ethnicity; Social Security Number; marital status; date of enrollment; referral status; expected source(s) of payment; charges; sources of informal care; living arrangement; physical limitations; functional impairment; diagnoses; number of visits; services received; and, for discharges, the outcomes of care. The facility and patient response rates in 1992 were 88 and 97 percent, respectively.

The next activity is the National Health Provider Inventory (NHPI). The NHPI incorporates and expands the NMFI which was conducted on a periodic basis since 1962 and was last conducted in 1986 as the Inventory of Long-Term Care Places. The NMFI served as an important source of national information on the inpatient health care facilities which provide medical, nursing, personal or custodial care. Sources and data collection methods used to compile these lists have varied over the years, but have included a mail questionnaire survey and an agency reporting system in which states and national organizations provide mailing lists. Periodically, complement surveys have been used to measure the undercoverage using area canvasses and searches of telephone directories.

The NHPI was conducted in 1991 to compile complete and current listings of nursing and personal care homes, home health agencies, hospices, homes for the mentally ill and mentally retarded, and licensed residential care facilities. The NHPI provides national statistics on these facilities and serves as a sampling frame for surveys of these settings. Data included name, address, ownership, types of services provided, number of beds, number and type of staff and patient data (annual discharges, annual admissions and current patients).

The NHPI will be conducted again in 1994 and will serve as the sampling frames for the 1995 National Nursing Home Survey and the 1995 National Home and Hospice Care Survey.

Research is also underway at NCHS to develop the NHCS Patient Follow-up Studies. The intent of these studies is to provide information on the outcomes and subsequent utilization of health care resources through periodic contacts with patients, their families or next of kin. By forming cohorts of patients based on demographic characteristics, diagnoses, treatments, mode of payment, etc., information may be collected directly either from the patient or the patient's family on subsequent treatment, outcomes of treatment, changes in health status and prior exposures to health risks. These data will be useful in assessing resource utilization and financing of health care, in evaluating the effectiveness of care, in tracking new technologies and treatments, in studying episodes of illness and the predisposing factors for seeking care.

An early example of such a longitudinal study was the follow-up study of the current and discharged residents sampled in the 1985 NNHS. Follow-ups were conducted in 1987, 1988 and 1990 to collect information on the resident's vital status, living arrangements, nursing home stays, hospital stays and sources of payment for stays.

More recently, in 1991, a contract was awarded to investigate the development of a general approach for conducting patient follow-up studies based on the NHCS. Research into the barriers to data collection and the use of personal identifiers of patients is being conducted. A field test of data collection methodologies for hospital discharges and hospital outpatient departments is being conducted in the Fall of 1993.

Summary

In 1989, the NCHS requested that the National Research Council through its Committee on National Statistics and the Board on Health Care Services of the Institute of Medicine convene a panel of experts to evaluate the plan for developing the NHCS. The panel's charge was to identify current and future needs for health care data and to determine the extent to which the NHCS could meet identified needs given the statistical aspects of the survey. In 1992, the final report from this evaluation was published and indicated a variety of broad recommendations and future directions for the NHCS.

The NHCS is an ambitious NCHS effort to provide needed information on the utilization of health care resources in the U.S. The NHCS is the result of a number of years of planning involving many Federal agencies and private organizations. With full implementation of this family of surveys, the NHCS will serve as a valuable resource for monitoring changes in the delivery of health care, the impact of health care reform and the effectiveness and quality of care provided to a changing U.S. population.

REFERENCES:

Bacon, WE. Health records surveys of the National Center for Health Statistics. Proceedings of the 11th International Health Records Congress. Vancouver, Canada: Canadian Health Record Association. 159-165. 1992.

Graves EJ. National Hospital Discharge Survey: Annual summary, 1988. National Center for Health Statistics. Vital Health Stat 13(106). 1991.

Haupt BJ, Kozak, LJ. Estimates from two survey designs: National Hospital Discharge Survey. National Center for Health Statistics. Vital Health Stat 13(111). 1992.

Hing E, Sekscenski E, Strahan G. The National Nursing Home Survey: 1985 summary for the United States. National Center for Health Statistics. Vital Health Stat 13(97). 1989.

Kovar, MG. Data systems of the National Center for Health Statistics. National Center for Health Statistics. Vital Health Stat 1(23). 1989.

Massey JT, Moore TF, Parsons VL, Tadros W. Design and estimation for the National Health Interview Survey, 1985-94. National Center for Health Statistics. Vital Health Stat 2(110). 1989.

National Archives and Records Administration. Federal Register 50(147). Washington: National Archives and Records Administration. 31038-40. 1985.

National Center for Health Statistics. Health, United States, 1991. Hyattsville, Maryland: Public Health Service. 1992.

Office of National Health Statistics, Office of the Actuary: National health expenditures, 1990. Health Care Financing Administration. Health Care Financing Review 13(1). 1991.

Roper, DA. Nursing and related care homes as reported from the 1982 National Master Facility Inventory Survey. National Center for Health Statistics. Vital Health Stat 14(32). 1986.

Schappert SM. National Ambulatory Medical Care Survey: 1989 summary. National Center for Health Statistics. Vital Health Stat 13(110). 1992.

Simmons WR, Schnack GA. Development of the design of the NCHS Hospital Discharge Survey. National Center for Health Statistics. Vital Health Stat 2(39). 1977.

Sirrocco, A. Nursing home characteristics: 1986 Inventory of Long-Term Care Places. National Center for Health Statistics. Vital Health Stat 14(33). 1989.

SMG Marketing Group, Inc. Hospital Market Database. Chicago, Illinois. Healthcare Information Specialists.

Strahan GW. Overview of home health and hospice care patients: Preliminary data from the 1992 National Home and Hospice Care Survey. National Center for Health Statistics. Advance data from vital and health statistics; no 235. 1993.

Tenney JB, White KL, Williamson JW. National Ambulatory Medical Care Survey: Background and methodology, United States. National Center for Health Statistics. Vital Health Stat 2(61). 1974.

Wunderlich, GS, ed. Toward a National Health Care Survey. Washington, D.C. National Academy Press. 1992.

THE OCCUPATIONAL SAFETY AND HEALTH SURVEY: INSTRUMENT REDESIGN

Leda Kydoniefs, Bureau of Labor Statistics
Bureau of Labor Statistics, 2 Massachusetts Ave., N.E., Room 3160, Washington, DC 20212

KEY WORDS: survey methodology, forms redesign, occupational injuries and illnesses.

Introduction

A major redesign of the Occupational Safety and Health Statistics (OSHS) system, the nation's primary source of information on job-related injuries and illnesses, began in 1988. The OSHS survey is a mandatory Bureau of Labor Statistics (BLS) survey, and is collected via mail in cooperation with State agencies. In order to fully convey to the reader the extensiveness of the OSHS redesign, this paper begins with a brief historical development and overview of the program.

Historical Overview[1]

Occupational Safety and Health Act

The Occupational Safety and Health Act of 1970 requires the Secretary of Labor to develop an effective program of occupational safety and health statistics. In 1971, the Secretary of Labor delegated to the Commissioner of BLS the responsibility of furthering the purposes of the Act "by developing and maintaining an effective program of collection, compilation, analysis and publication of occupational safety and health statistics." The Secretary further directed the Commissioner of BLS to coordinate the above functions with the Assistant Secretary for Occupational Safety and Health.

OSHA Recordkeeping System

The recordkeeping system, which is the foundation of BLS's statistical program, was developed to aid the Occupational Safety and Health Administration (OSHA) in setting standards, to assist safety and health officers in identifying hazardous operations, and to provide BLS and State agencies with uniform and reliable safety and health statistics. Further, the statistical program would provide employers and employees with information about conditions at their workplace.

In 1978, a recordkeeping form designed to streamline OSHA mandated recordkeeping and reporting was implemented. It is referred to as the Log and Summary of Occupational Injuries and Illnesses (OSHA 200) or simply the OSHA Log. Until 1991, BLS managed the Log for OSHA, distributing copies to employers, and ensuring that recordkeeping guidelines were conveyed clearly. The employer is mandated to record onto the OSHA Log all work-related deaths, illnesses, and injuries. Furthermore, the employer is required to describe each case in detail using a supplementary record such as the OSHA 101 provided by OSHA, or by using some equivalent form, (i.e., a form that includes all necessary information) such as a State worker's compensation form. Distribution of OSHA 101 was also managed by BLS for OSHA until 1991.

Annual Survey of Occupational Illnesses and Injuries

Shortly after the Occupational Safety and Health Act was enacted, BLS in cooperation with State agencies began to conduct the Annual Survey of Occupational Injuries and Illnesses providing national measures of the incidence and severity of workplace injuries and illnesses. The Annual Survey is collected via mail from a probability sample of approximately 280,000 establishments in private industry. Participation in the survey is mandatory. Establishments with 11 employees or more are in scope, and self-employed persons are not covered by the survey. Prior to the redesign, each sampled establishment provided summary totals of the number of injury, illness, or fatality cases experienced during the calendar year. The total numbers of lost work days and days of restricted activity were also collected, as was annual average employment and total hours worked by employees. Employers copied information for the injury and illness portion of the Annual Survey directly from the OSHA Log that they maintained throughout the year. All employers covered by the Act are required to maintain the Log with the exception of employers with fewer than 11 employees, and employers in low-risk industries such as retail and real estate. The Annual Survey samples employers who are ordinarily exempt from this recordkeeping requirement, but these sampled employers are required to maintain the OSHA Log for the year they are in the survey.

Supplementary Data System

Not long after the Annual Survey began, safety and health analysts and other interested parties identified a need for greater detail than that of incidence rates. They requested information about the characteristics of the occupational injuries and illnesses and the workers to whom they were occurring. It was generally recognized that records routinely generated by

---

[1] Portions of the Historical Overview were taken from the BLS Handbook of Methods, Bulletin 2285.

State workers' compensation programs such as employee and employer reports, medical reports, compensation award records were all valuable sources of information about occupational injuries and illnesses. As a result, State workers' compensation agencies began, in 1976, to participate in BLS's Supplementary Data System (SDS).

The SDS program was conceptualized as providing data to "supplement" that provided by the Annual Survey. The program's source of information was a first report of injury or illness, which employers and insurance carriers submit to State workers' compensation agencies. SDS data have helped in the identification of general patterns in the characteristics of work-related injuries and illnesses. The SDS program, never designed as a nationwide system, was established in cooperation with 27 States, and included anywhere from 12 to 36 cooperating States in a given survey year. Although the SDS provided valuable information and standardized the classification, processing, and tabulations of data, it was never a complete census of occupational injuries and illnesses. In addition, it suffered from variations in coverage and reporting requirements that reflected differences in State workers' compensation laws.

Programmatic Review

By the mid-1980's OSHA, BLS and numerous representatives of business and labor had begun to feel that the existing BLS safety and health surveys should be integrated into one internally consistent statistical system. For example, the National Academy of Sciences (Pollack, & Keimig, 1987) in reviewing BLS programs felt that the Annual Survey did not provide needed information about the worker and the circumstance of the case, within an industry context. Also, the SDS was deemed insufficient because of its partial case coverage and lack of consistency across all States. It was recommended that the Annual Survey be modified to include specific categories of injuries and illnesses, thereby folding SDS type of data and incidence data into one internally consistent program that could produce national estimates on both incidence rates as well as case characteristics.

The BLS, in responding to these recommendations began an all-inclusive redesign of the Occupational Safety and Health System (OSHS). The redesign effort has included a review of all aspects of OSHS: Survey content, sampling, forms design, collection and processing. This paper will focus on the survey instrument redesign effort: changes in content and question wording; forms redesign; methodology utilized to test the "new" survey format; the data analysis and resulting survey revisions.

Revisions to Survey Content

The OSHS survey was expanded to include details about the specific injuries and illnesses incurred by employees. The following details were added to the survey:

Details About the Injured Worker

Name and social security number
Gender, Age, and Ethnicity
Length of service at the establishment
Occupation

Details About the Injury or Illness

When did the injury occur?
How many days of lost work or days of restricted work activity resulted?
What was the employee doing just before the incident occurred?
What happened - how did the injury or illness occur?
What was the injury or illness - i.e., what part of body was affected and how was it affected?
What object/ substance directly harmed the employee?

The employer is asked to provide the above information on each illness or injury case that occurs in the workplace. A prediction is made as to the expected number of cases an employer will have, and the appropriate number of "case forms" are provided. If the employer is expected to have more than 20 cases, instructions are included in the survey directing the respondent to sample from their list of injuries and illnesses so as to minimize, as much as possible, their response burden .

Survey Redesign: Research and Development

Redesign Overview

The survey redesign portion of the project began in 1989 and consisted of several pilot/feasibility studies investigating alternative survey formats. As previously discussed, programmatic emphasis was placed on expanding the data base to include not only incidence rates of injury and illness by industry, but to provide demographic characteristics of the injured worker as well as begin to collect detailed information on the incidence itself. In 1991, the decision regarding the expanded scope of the survey was finalized. The complexity of the survey was such that the design of a booklet format became a necessity, and forms redesign efforts intensified. The objectives were four-fold:

1. Develop a booklet survey form with accompanying instructions that was "user friendly."
2. Design a questionnaire format with a built-in flexibility that would permit certain minor vari-

ations in item content. (These variations are required in order to maximize the survey's utility across the nation.)

3. Test and evaluate old (i.e., rephrased) questionnaire content as well as new survey items.
4. Test and evaluate the final product, i.e., the redesigned OSHS survey that was developed.

To achieve these goals, a variety of behavioral science and test development methods were used, each of which added a different perspective to our knowledge base and forms redesign efforts. (For an overview of these methods see Forsyth & Lessler, 1990; Gower & Nargunkar, 1991; and Jabine, Straf, Tanur, & Tourangeau, 1984.) Focus groups (Krueger, 1988) were used primarily in the beginning of our test development efforts to refine the scope of the survey, the item wording, and obtain a gross estimate of respondent burden. Respondent observations using a think-aloud protocol (Mullin, Miller, Melis-Wright, & Stone, 1981) were used when the survey redesign approached completion. Pretest mailing of the survey (Dillman, 1991) with an accompanying respondent (probing) questionnaire served as the last step in the "test development" phase prior to fielding the survey. These "pretest" respondents were also contacted via phone and asked additional questions regarding their response to the redesign of the survey.

The sections that follow present the developmental steps that led to the final survey booklet: the early feasibility studies, pilot studies, and repeated testing of draft survey booklets.

Feasibility Studies

During 1988, twelve feasibility studies (Bureau of Labor Statistics, 1988) were conducted by BLS. Cumulatively, the studies collected data from 20 States and sampled about 100 establishments per State (a total sample of 2000 establishments). These studies were exploratory in nature designed to find out what kind of records/data employers kept on workplace injuries and illnesses, and to inquire whether employers were willing to share these data with BLS.

The research was conducted via mail, and respondents were asked to voluntarily submit copies of particular records. Depending on the study, employers were asked to send in copies of the OSHA Log, first reports of injury, and/or injury forms filed with Workers' Compensation. The feasibility studies yielded encouraging results: Most employers maintained the OSHA Log and some accompanying first report of injury. Moreover, they were generally willing to report on the data contained in these documents.

As a result of these positive findings, BLS proceeded to further elaborate and clarify the data elements that would eventually be added to the Annual survey. To assist in their effort, BLS solicited input from the survey data user community requesting suggestions for data elements of interest to various groups. A variety of organizations were approached: OSHA, National Institute for Occupational Safety and Health (NIOSH), the BLS Labor Research Activity Council (LRAC), the BLS Business Research Activity Council (BRAC) as well as many State Labor and State Health Departments.

Pilot Studies: Phase I

In 1989, BLS returned to the field for additional research. The primary objective was to investigate which of the numerous data elements generated internally by BLS staff and/or suggested by the survey user community were collectible. A second objective was to investigate the utility of certain basic changes in existing records and/or data collection formats. The reader will recall that in addition to conducting the Annual Survey, BLS managed (for OSHA) the OSHA Log and OSHA 101 - the supplementary record developed to provide case injury or illness detail. With these pilot studies BLS was exploring improvements to all safety and health related data collection formats.

Five pilot studies were conducted in a total of fifteen (13) States (Bureau of Labor Statistics, 1989). About four hundred (400) establishments were sampled within each State yielding a total sample of approximately 5200 establishments, or 1040 establi-shments per pilot study. The five pilot studies were as follows:

Pilot 1. Testing a record keeping format that combined the information collected via the OSHA Log and first report of injury (e.g., OSHA 101) into one form. The employer was also asked to code cases identifying the nature of the injury, the part of body effected, and the source of the injury.

Pilot 2. Testing the same record keeping format described in Pilot 1 above, with the exception that the employer did NOT have to precode data.

Pilot 3. Testing a revised supplementary record (OSHA 101). As with Pilot 1, the employer was asked to precode case characteristics.

Pilot 4. Testing the revised supplementary record described in Pilot 3 without the coding of case characteristics.

Pilot 5. Within this study, respondents were asked to submit a copy of the OSHA Log. The State then attempted to match the injury and illness cases identified on the log to the State workers compensation files submitted by the establishment, and extract the appropriate data elements regarding case characteristic details. The reader will recall earlier discussion (see Historical overview) on the detailed data generally available within State workers' compensation agencies.

Analysis of the five pilot studies resulted in a greatly reduced list of data elements under consideration for inclusion in the Annual Survey. Further, it was concluded that employer self-coding was not a feasible pursuit. Employers made too many errors when asked to apply an unfamiliar coding scheme to their own records.

Pilot Studies: Phase II

In 1990, BLS began a second phase of pilot studies. Starting from where the first set of pilots left off and applying the knowledge gained, BLS proceeded to refine and retest the collection methodologies described in Pilots 2, 4, and 5 above. As before, each pilot study was conducted in three States (9 States in total) and approximately 400 establishments per State, yielding a sample size of approximately 1200 establishments per pilot study. Once again, employers were asked to:

a. Complete a record keeping format that combined the information collected via the OSHA Log and first report of injury into one form.
b. Complete a revised supplementary record (OSHA 101).
c. Submit copies of the OSHA Log with the State following improved procedures to match cases identified by the employer to the workers compensations forms filed with the State.

As this phase of pilot studies was being conducted, OSHA and BLS agreed to separate the mandatory record keeping requirements and accompanying documents from the statistical survey or data collection aspects of occupational illnesses and injuries. This separation was welcomed by BLS, interested in increasing the distance between its pledge of confidentiality and data collection for statistical purposes only and OSHA - a government agency responsible for assessing fines for noncompliance to Safety and Health regulations. Consequently, BLS transferred all responsibility for the distribution and revisions to the OSHA Log and the supplementary record (OSHA 101) to OSHA. Taking the data collected thus far from the Phase II pilot studies, as well as the earlier BLS feasibility and Phase I pilot studies, OSHA began its own efforts at revisions to the OSHA Log and accompanying supplemental record. BLS was free to focus solely on implementing revisions to the statistical portion of occupational injuries and illnesses program, i.e., the Annual Survey. Momentum increased as development work was directed solely on the survey booklet - "Survey of Occupational Injuries and Illnesses".

Development of Survey Booklet

As the pilot studies were concluded, BLS entered a different stage in the survey instrument redesign. BLS now had a clear sense of direction with respect to the type of questions that would be added to the survey, and felt confident regarding the availability of employer records on the data of interest. It was decided that in order to accommodate the additional data elements and provide (survey completion) instructions that were clear and easy to follow, the survey instrument would have to be re-conceptualized. The Annual Survey thus changed from a 6 x 11 multi-part pin fed form, to a 8.5 x 11 multipage booklet. Development work from this point on focused on testing and refining the booklet format, the wording of newly developed items, and the accompanying survey instructions. Such testing required a different methodology from the large scale pilot testing conducted thus far. Sampling a large number of respondents at a time was no longer necessary. What was needed was to access a few respondents at a time and ask them to look at the survey booklet and provide BLS with detailed feedback.

First Draft: Focus Group and Respondent Interviews

In September 1990, a study (Palmisano, 1990) was designed to: (a) investigate the clarity of newly developed item wording and (b) evaluate the respondent's perceptions of a test version of the survey booklet -- how the content and format would impact on the ability to effectively, efficiently and accurately complete the survey. The study was conducted in two phases: the first consisted of a group of twelve (12) BLS employees, and the second consisted of five (5) individual interviews with representatives of business establishments from the private sector. All interviews were videotaped to assist with later analysis. BLS employees were chosen, partially for ease of access, but primarily for their lack of knowledge of safety and health related forms and recordkeeping guidelines and for their extensive experience with various other "government forms". It was felt that this latter group of individuals would be in a position to comment on the strengths and weaknesses of a BLS survey having as a comparison the various forms they worked with on a daily basis. The five representatives of business establishments had experience completing the OSHA recordkeeping forms as well as the Annual Survey, and could comment on how this "new" booklet faired in comparison.

Furthermore, the two phases were expected to yield very different kinds of information: From a focus group, one expects the feedback/information to grow out of the group interaction -- one group member adding to, confirming, or disagreeing with that said by others. With one-to-one interviews, the focus is solely on the interaction between the respondent and the instrument. From such observation, one can expect greater emphasis on details and a more carefully considered reaction from the respondent. Being alone, the respondent has to feel strongly enough about a comment in order to verbalize it and cannot, as in a group, simply agree with what was stated by others.

Respondents were first asked to complete the survey booklet, and then to discuss the document listing its best and worst features, and why they thought particular features were important. The survey booklet consisted of the following: a section asking for summary data of injuries and illnesses; a sampling matrix designed to enable the employer with many cases (of injuries and illnesses) to sample a smaller number of those cases and thus reduce response burden; and copies of a case form designed to collect individual details from the sampled cases.

Data analysis revealed weaknesses in the clarity of instructions and certain item wording but, most importantly, revealed problems with the sampling matrix. Specifically, respondents found the case sampling matrix overly 'mathematical' and difficult to use, and the accompanying instructions confusing. Additionally, differences between the BLS employees and representatives of business in the type of difficulties encountered with the booklet revealed invaluable information about a "private citizen's" attention span in reading "government instructions". Instructions would have to become greatly simplified to ensure that respondents would understand them after only a cursory glance.

Results of this study led to a thorough review and revision of the contents and form of the survey booklet.

Second Draft: Respondent Interviews and Mailout Testing

BLS returned to the field in 1991 to conduct further field tests. The revised booklet was subjected to two different types of testing: intensive "cognitive" testing involving interviews with one respondent at a time, and mass mailing of the booklet (Palmisano, 1991). Each methodology was designed to test a different aspect of the booklet.

Second Draft: Interviews with Respondents

It was decided to interview respondents from a broad range of geographic locations, establishment size, and industrial classification. Nineteen (19) interviews were conducted within establishments located in four States. Half of the establishments sampled were members of industries with typically high rates of recordable injuries and illnesses, and the remaining half were members of industries with low rates of recordable injuries and illnesses. Further, half of the sample consisted of establishments with more than 50 employees; the other half consisted of less than 50 employees. As may be obvious, establishment size and typical incidence rate for the establishment industry speaks to the level of sophistication and experience we expected a respondent to bring to bear in reviewing and commenting on the test survey booklet.

Each respondent was interviewed (and videotaped) within his/her establishment. As with the study described above (see First Draft) respondents were given a survey booklet containing a section asking for summary data of injuries and illnesses; sampling instructions designed to enable an employer with many cases (of injuries and illnesses) to sample a smaller number of those cases; and copies of a case form designed to collect individual details from the sampled cases. This time the respondents were also given a mock OSHA Log listing several "cases" and a number of completed mock supplementary forms. The latter contained the detailed characteristics pertaining to the cases listed in the mock Log. The supplementary forms were workers' compensation forms appropriate to each State in which the testing took place. These mock documents were provided to respondents in order to achieve standardization of recordkeeping documents across respondents. The focus of the study was on the interaction between each respondent and the survey booklet. Standardizing the documents respondents needed to use in completing the test survey booklet was quite important as it prevented the results from being confounded due to any variability of records maintained across establishments. A great deal of time was invested in generating these mock documents. It was deemed important that the documents be representative of what is typically found within establishments, so care was taken to ensure that a variety of injuries and illnesses were represented; not all the data were complete; the records were not typed; and handwriting was not always legible, etc.

Respondents were asked to complete the test survey booklet using the Log and supplementary forms provided. They were not allowed to ask for assistance while they were in the process of completing the booklet. Respondents were told in advance that they should note any problems encountered directly on the booklet and, despite the problems, should continue as best they could. After they completed the survey, all respondents participated in a debriefing interview focused on elaboration of the problems experienced while completing the survey.

Second Draft: Large Scale Mailing

Survey booklets were mailed to a total of 1275 establishments located in three States. The test sampled establishments who, based on previous survey years and knowledge of industrial classification were expected to have a large number of cases to report . Completion of the test booklet was voluntary, as indeed was participation in all the tests discussed in this paper. Employers were informed via an enclosed letter that they had been selected to take part in the testing of the OSHS redesign, and were requested to complete the survey booklet.

Analysis of results (Finch, Enochs, Carder, 1991; Laundrie, 1991; Winter, Young, & Dorton, 1991) focused not only on the quality of the completed surveys, but also on an evaluation of the data collection and processing aspect of this newly developed survey format. In other words, it was important not only to develop a survey booklet that was easy for a respondent to understand and complete, but the booklet had to lend itself to internal (BLS) review and processing. The new booklet had to pass two litmus tests if you will: one by respondents, and the other by the cooperating States involved in processing the data.

Data from both the large scale mailing and the interviews with individual respondents were analyzed and the booklet was reviewed extensively. It was concluded that the wording of the new data elements was acceptable, as was the wording of most of the instructions. However, the forms design needed to be improved and simplified and, more importantly, so did the sampling matrix. Respondents were still having difficulties knowing which cases they were required to report on the survey booklet -- the instructions for sampling from the cases listed on their OSHA Log needed to be simplified. Appropriate revisions to the survey booklet were implemented and BLS returned to the field for another round of testing.

This particular set of revisions was pivotal in the OSHS program redesign process. Data had been accumilating that the survey data collection instrument needed to be simplified as much as possible. Further, BLS staff were becoming increasingly concerned that the size of the booklet (as conceived at this stage of development) would prove too daunting to many respondents and thus reduce response rate. As a result, it was decided to produce four different versions of the survey booklet, each version containing a different number of copies of the page (within the survey booklet) designed to collect details about the individual injury and/or illness. Based on prior years of survey data collection, BLS is able to make a prediction regarding the number of cases with injury and illness each sampled establishment is expected to have. Establishments expected to have a large number of cases with injury and illnesses would be mailed a survey booklet with a larger number of these "case pages", and establishments expected to have a few cases, would be mailed a booklet containing a few copies of the case page. Obviously, producing more than one version of the survey booklet greatly complicated the survey printing, mailing and quality control procedures as well as the overall cost of the OSHS survey program. Nevertheless, this decision was in keeping with programmatic mandate for the OSHS redesign -- to develop a "user friendly" instrument while collecting more detailed information on injuries and illnesses.

Throughout the development of the various drafts of the survey booklet, the BLS benefited from the assistance and feedback of the State Advisory Committee on Occupational Safety and Health (SACOSH.) This advisory committee was chartered to advise the BLS on matters related to OSHS redesign. It is composed of members of the States involved in collecting the OSHS data. Certain SACOSH members were proponents of developing a survey booklet that would request minimal data from the respondent and require the State to collect the rest of the data by accessing the State Workers Compensation data files. For States having the facility to access the workers compensation files via computer, this was seen as a viable option and a further move towards reducing respondent burden. Consequently, BLS produced yet another variation of the survey booklet: The "workers compensation survey booklet". It looked very similar to the "standard booklet", and also had four versions each containing different numbers of the individual case page. However, the "Workers compensation" booklet asked the respondent at a certain point within the individual case page to proceed no further if a particular case had been submitted to the State workers' compensation office, thus effectively reducing the time needed to complete the survey.

Third Draft: Think-Aloud Protocol

In early 1992, BLS began the final round of tests on the survey booklet (American Institutes for Research, 1992). Revisions as a result of information learned during second draft testing (see above) had produced a survey booklet that BLS staff felt comfortable with. The survey booklet instructions were clear and simple to follow. The forms design was utilitarian, providing visual clues early on in the document that assisted the respondent in maneuvering through the form. The sampling method (for employers expected to have many cases of injuries and illnesses) was greatly simplified and could be followed without a great deal of effort.

Nine (9) establishments were sampled and individuals working in the safety or health departments interviewed and videotaped. The respondents completed the test booklet using a think-aloud protocol and their own injury/illness records as source documents. The respondents were asked to complete the test booklet one page at a time and encouraged to "think aloud" or verbalize how they interpreted the material they were reading, and what questions they had along the way. At the end of each page of the booklet, respondents were asked prepared questions and explored any problems they had along the way. At the end of the test session, respondents were asked to complete two questionnaires: one about the test booklet and the other about their background and familiarity with Safety and Health recordkeeping documents.

Analysis of the videotapes revealed very positive findings. None of the participants had problems with the overall design of the survey, the case sampling, nor the majority of instructions. These instructions were revised as were a few items found to be in need of minor wording changes, and reordering.

Final Draft: Pretest Mailout and Respondent Debriefing

The final draft of the survey was mailed to 215 respondents located in six States (American Institutes for Research, 1992). As with the second draft of the booklet, respondents from a broad range of establishments in terms of size and industrial classification were sampled. The sampling was purposeful, designed to obtain respondents with maximum variability in exposure to safety and health regulations and cases of work-related injuries and illnesses. Each respondent was mailed a test survey booklet, a survey questionnaire, and a cover letter describing the purpose of the test. The survey questionnaire asked respondents questions about their experience in completing the booklet: the time it took to complete the survey; availability of in-house records needed as source documents; various probes related to answers provided; suggestions for improvement, etc.

After completed survey booklets were received and reviewed, respondents were called and asked follow-up questions about the survey. This "respondent debriefing" included prepared questions asked of all respondents as well as specific questions related to the way a particular respondent completed the booklet. It is interesting to note that all establishments sampled had been in the sample for the Annual Survey that year, and therefore, had completed the "official survey" a few months prior to this mailout test. Despite this additional burden all respondents were obliging and giving of their time.

Analyzing the completed survey booklets, the questionnaires, and evaluating the phone interviews, it was concluded once again that respondents had little trouble with the overall design of the survey. There were very few problems with the survey's wording, and areas where some employers became confused were minor enough so as to warrant minimal changes to the final survey booklet. Phone interviews with respondents proved particularly useful in determining the sensitivity respondents had to certain items. Though the sampled respondents completed all items, they did not particularly like providing certain information. Respondent debriefings are particularly useful in eliciting information that cannot not be obtained from a review of the completed survey booklets. The importance of such information is further underscored when one considers that in a survey with a sample size of 280,000 establishments (such as the OSHS) such response sensitivity could reflect thousands of complaints and perhaps even a decrease in response rate. In this particular case, the sensitive items were deemed to be of great importance to the survey and were retained. However, other steps were taken (programmatically) to reduce any potential negative impact on the survey and its contents.

Although specific details of any portion of the OSHS redesign are beyond the scope of this paper, two examples of steps taken to reduce potential negative impact of the redesign should be mentioned: The BLS engaged in an extensive "outreach" effort that involved the development of outreach materials (slides, outlines, talking points) for use by the cooperating State agencies. An outreach brochure was developed to announce the arrival of the redesigned Annual Survey, and was sent to establishments that had been in the sample for more than one year. Furthermore, numerous national organizations were contacted via mail and asked to include a notice of information about the Survey changes in mailings to their constituents. Many organizations included this notice in their newsletters and/or magazines. Others chose to send the notice under separate cover to their membership.

An alternate method developed to reduce any potential negative impact of the survey, its changes, and sensitivity to its contents was providing an entire page in the survey booklet for respondent comments and feedback. The backpage of the survey booklet, designed to appear friendly and inviting, requests comments or suggestions about the survey -- what respondents liked as well what they didn't. The BLS is sincere when adding: "We will review and carefully consider your comments and ideas as we work to improve our survey." BLS is currently in the process of collecting, collating and coding all the comments we have received. Soon, we will be in a position to address revisions to the "redesigned" Annual Survey of Occupational Injuries and Illnesses. The process will begin all over again and we will return to the field to test, once more, changes we have made to the survey booklet. In light of lessons learned these past few years of redesign, the OSHS survey is now considered to be an ever changing program under a process of continual improvement.

REFERENCES

American Institutes for Research. (1992, September). Redesigning and Testing the 1992 Survey of Occupational Injuries and Illnesses: Final Report. (Contract No. J9-J1-0067) Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.

Bureau of Labor Statistics. (1988). BLS Handbook of Methods. (Bulletin 2285). Washington, DC: U.S. Government Printing Office.

Bureau of Labor Statistics. (1988). Occupational Safety and Health Statistics Program: 1988 Feasibility Studies. Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.

Bureau of Labor Statistics (1989). Report on the 1989 Occupational Injury and Illness Pilot Surveys. Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.

Dillman, D. (1991) The Design and Administration of Mail Surveys. Annual Reviews of Sociology, 17: 225-249.

Finch, B., Enochs, D., & Carder, S. (1991, October). Oklahoma Pilot "C" Study: Final Report. Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.

Forsyth, B. H., & Lessler, J. T. (1990, November). Cognitive Laboratory Methods: A Taxonomy. Paper presented at the International Conference on Measurement Errors in Surveys, Tucson, AZ.

Gower, A. R., & Nargunkar, M. S. (1991). Cognitive Aspects of Questionnaire Design: Business Surveys versus Household Surveys. Paper presented at the Annual Research Conference, United States Bureau of the Census, Washington, DC.

Jabine, T. B., Straf, M. L., Tanur, J. M., & Tourangeau, R. (1984). Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines. Washington DC: National Academy Press.

Krueger, R., A. (1988). Focus Groups: A Practical Guide for Applied Research. Newbury Park, CA: Sage Publications.

Laundrie, S. P. (1991, October). Maine Pilot "C" Study: Final Report. Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.

Mullin, P., Miller, L. A., Melis-Wright, M., Stone, D., (1991, April). Laboratory Research Methods for Improving Surveys. Paper presented at the GAO Technical Conference, Washington, DC.

Palmisano, M. (1991, November). The Evaluation of a Refined Occupational Injuries and Illnesses Questionnaire. Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.

Palmisano, M. (1990, October). Respondent Reactions to a Proposed BLS Survey Questionnaire for Occupational Injuries and Illnesses. Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.

Pollack, E. S., & Keimig, D. G. (Eds.). (1987). Counting Injuries and Illnesses in the Workplace: Proposals for a Better System Washington, DC: National Academy Press.

Winter, D., Young, R., & Dorton, H. (1991, October). Georgia Pilot "C" Study: Final Report. Washington, DC: Bureau of Labor Statistics, Office of Compensation and Working Conditions.