RESPONSE ERROR RESEARCH ISSUES IN HEALTH SURVEYS

D. G. Horvitz, Research Triangle Institute

I. Introduction

The papers in this session address several statistical aspects of the use in medical care surveys of data available in the records maintained by medical care providers, including hospitals, clinics and office-based physicians. Particular attention is given in three of the papers to evaluation of procedures for combining the medical provider record data (referred to as record check data) with the survey reported data. The process of combining medical care survey data with useful information from the records maintained by medical providers requires matching the two data sources at analysis unit levels; for example, hospital admissions or physician office visits recorded or reported during a specified time period. The matching process is subject to error and the effects of matching errors will also be discussed in the last paper of this session.

The purpose of this paper is to provide some background for the subsequent papers and to raise related survey research issues some of which are addressed in this session, at least in part, by the other speakers. These issues have arisen in conjunction with two recent panel surveys of the use of medical care services in the United States and the related expenditures for the care received. The first of these surveys, the National Medical Care Expenditure Survey (NMCES) sponsored by the National Center for Health Services Research (NCHSR), with support from the National Center for Health Statistics (NCHS), covered health care utilization and expenditures by the U.S. non-institutional population during calendar year 1977. The second survey, designated the National Medical Care Utilization and Expenditure Survey (NMCUES), provides similar comprehensive data for the U.S. non-institutional population for 1980 and for Medicaid eligible families in four states, California, Michigan, New York and Texas. NMCUES is jointly sponsored by NCHS and the Health Care Financing Administration (HCFA).

Both NMCES and NMCUES are panel surveys in the sense that the data are collected for the year of interest by a series of periodic interviews with the initial sample of households. The principal purpose of the repeated interviews at intervals of 13 weeks for these surveys was to improve the quality of the data. In both surveys, the key data items include the details of each dental, doctor, clinic or emergency room visit and each hospital stay, including dates and services received; the charges for the health care services received; prescribed medicines purchased and their costs; other medical expenses and finally the source of payment for the care received, that is, how much was paid out-of-pocket by the family, and how much was paid by an insuror or other third party, whether public or private.

Despite the use of repeated, bounded interviews every 13 weeks along with a calendar/diary and a computer-generated Summary of the key data for each previously reported medical care visit prepared and mailed to the respondents prior to each interview, the quality of the NMCES respondent reported data on utilization, diagnoses and expenditures was still considered sufficiently suspect as to require a record check survey of those physicians, clinics, hospitals and other medical providers who had supplied health care to a subsample of the NMCES participants. The record check survey, designated the Medical Provider Survey (MPS), is expected to fill in gaps in the NMCES data, particularly with respect to unreported visits, diagnoses, total charges and sources of payment, as well as to provide an opportunity to adjust for inaccuracies in the family and individual respondent reported data. Although a similar medical provider survey was not included as a part of NMCUES, record check claims data will be obtained for Medicare respondents in the national survey and for Medicaid respondents in each of the four state surveys. The use of the Medicare and Medicaid claims data in combination with the respondent reported data is particularly appropriate since the quality of health survey data is somewhat less for the elderly and the poor than for those younger and better off economically (Andersen et al.).

II. The Role of Survey Error

Users of sample survey data may be classified into two major categories. There are those who ignore the existence of any error in the data, a category that has been growing smaller, fortunately, year by year; the second, and increasing category of users insists on knowing something about the level of error in the data in order to have better control over its own errors of interpretation. A third, very minor and unfortunate, category can hardly tolerate error of any kind in survey data. Those who design surveys form a fourth category, also minor, but clearly important. They must not only be aware of each and every source of error, they must devise techniques which permit cost effective control of the level of total error in estimates to be derived from survey data. In a sense, they must, like the engineer, make error work for them. I refer to quality control mechanisms which measure the deviation, or the level of error, from desired system operating characteristics and, based on the magnitude of the error measurement, generate feedback signals which set corrective machinery in motion. The key, of course, is knowledge of the level of error. For example, without some knowledge of the relationship of bias to response rate one cannot sensibly allocate survey resources as between additional sample units and followup, or other special efforts, to reduce nonresponse.

A major difference between the engineer and the survey statistician is the opportunity to measure error levels on-line, that is, concurrent with system operations. Certain data collection operations are very amenable to on-line quality control mechanisms including interview validation, field supervisor edit and central office edit of interviewer's completed questionnaires. Data processing operations such as editing, coding and data entry are also readily subject to on-line quality control. I suspect that nonresponse

bias can also be addressed on-line, at least partially. The opportunity to detect and hence correct response errors during the course of the data collection phase of a survey is limited usually to gross errors detectable by error checks in computer-assisted telephone interviews or in the edit process.[1/] The accessibility of validating records is rarely rapid enough to permit on-line quality control of response errors, although longitudinal panel surveys offer some limited potential to do so. For example, Medicare numbers reported by NMCUES respondents during the initial interview were verified against national files during the course of the survey, permitting followup in later interview rounds of respondents whose numbers failed verification.

In the absence of adequate on-line quality control of response or measurement errors, techniques which rely on the use of record check data will still be necessary for variables which survey respondents are unable to report accurately. The record check data collected in a given survey can be used to assess the response error bias and the response error variance for respondent reported items available in the validating records. They can also be used in combination with the given survey data to produce more accurate estimates than would be possible from the survey data alone. Of course, the mean square error of the estimated bias must be less than the square of the bias for an adjusted estimate to be better than the unadjusted survey estimate. Finally, the estimated response error biases and response error variances derived from record check data for a particular survey constitute a very important resource for the design of subsequent surveys measuring the same variables.

III. Use of Administrative Records

Medical records, medical care claims records and administrative records maintained at the patient level are all sources of record check data for health care surveys. Such records have always been viewed as important sources of information and numerous studies have relied on them exclusively. They have also been used in reverse record check investigations of response errors in personal interview surveys concerned with usage of medical care facilities.

Medical and administrative records by themselves have not been used extensively in research requiring comparisons of population groups primarily because the size and characteristics of the populations represented by the records are not known with sufficient accuracy. They also suffer by their lack of detailed demographic and socio-economic measures which are often very essential to population based research. Finally, administrative records are not designed and maintained to meet survey research needs and suffer from unknown levels of error in the items recorded as well as lack of completeness. Marquis contends that errors in hospital records are partially responsible for apparent overestimates of hospital admissions found in forward record checks of survey data.

Interest in the use of medical and other administrative records to assess the quality of survey data has increased in recent years, due in part to the increased availability of machine readable files, but also due to a felt need for more complete and accurate information than are reported by survey respondents. The quality of the 1980 Census is being assessed in part by matching a 1978 Current Population Survey sample to Internal Revenue Service income tax files and to Medicare enrollee files using Social Security number. The entire process is rather complex and has a special focus, but considerable value should still be derived from this exercise by survey practitioners in general.

The design of health surveys which combine standard household samples with record checks for respondents reporting one or more visits to medical providers offers some challenge. This problem has been addressed in interesting fashion by Folsom et al., but needs further attention. The key to more optimal distribution of the survey resources between the household sample and the record check sample depends on the appropriateness of the error and cost models. The error model should recognize the differential accuracy with which details of visits to medical providers are reported by type of provider and by type of respondent (Andersen et al.). The error and cost models may also be different depending on the parameters to be estimated and the estimation procedure to be used. For example, are only means and totals to be estimated or are distributions, correlations and regressions also to be estimated? Are the record check data and survey data to be combined using a regression estimator, dual record estimation, imputation, weight adjustments or Bayesian techniques? Further research is clearly needed to determine the extent to which the optimum size and distribution of the household and record check samples vary according to the parameters to be estimated and the method of estimation.

IV. Matching

Optimum use of record check data in health surveys requires matching of survey reported medical care visits to the visits recorded for the same individual in medical provider or claims files. Visit level matching may not always be feasible and the analyst may have to be satisfied with matching at the patient level by type of provider. Regardless of the level of matching, there will be matching errors. Visits that should be matched will not, visits that should not be matched will. There are three research issues of interest in this context which arise because of matching errors. First, how are the alternative methods of combining health survey and record check data affected by matching errors; second, how should the error models used to allocate resources between survey samples and record check samples be modified to include matching error components; and third, what variables and matching criteria, including tolerance levels, will minimize errors in matching survey reported medical care visits with medical provider visit records for the same individuals.

1/ It may not be too long before interviewers in face to face interviews will use hand-held microcomputers which will control errors in recording the data, skip patterns and response inconsistencies.

V. Dual Systems

For the most part, the discussion thus far has implied the use of record check data to validate survey reported data. This assumes the medical and administrative records are sufficiently accurate to ignore any component of error due to faulty records. This is not exactly the case with medical care records and probably with other types of records as well. Some assessment is essential of the consequences of assuming the record check data on medical care visits to be free of error versus a combined or dual estimation procedure which assumes errors in both data sets.

A dual systems approach generates design issues. It suggests selecting a sample of individuals and independently generating two medical care histories for each of them, the first by a personal interview, the second by a search of medical provider records. This seems hardly feasible, even with signed permission from each person in the sample, since the number of eligible medical providers and hence, separate record files, is very large. Dual systems involving an interview survey and administrative records are clearly most feasible when there is a single set of records in a central file or at worst a small number of separately maintained files to be searched.

An advantage of the dual system approach is that it provides an opportunity to estimate the number of visits to medical care providers not reported by either data system. However, there is some reluctance to accept as real those visits reported by respondents to specific providers for which there is no record in the files of the medical provider named by the survey respondent. Some effort to determine the level of error in reporting the name of the medical provider by survey respondents seems essential to either verify, or provide evidence contrary to, this reluctance. This again implies searching the records of a large number of potentially eligible medical providers and may not be feasible. As an alternative, survey respondents could be asked to name all other medical providers they have ever visited in addition to those reported for the time period of interest. The files of the additionally named providers could then be searched for visits by the particular respondent during the time period of interest.

The net difference between medical care visit statistics as reported by interview respondents in a survey and those available in the files of the medical care services community might be estimated by selecting independent samples, one a sample of individuals for the survey and the second a sample of medical providers. Matching would not be feasible and the opportunity to estimate gross errors at either macro or micro levels would be lost. On the other hand, it may be possible to select a sample of patients from each medical provider in the first stage sample and interview them for the socio-economic data essential for a breakdown of the medical care visit statistics by population group. Again only net differences can be estimated.

This latter design points in the direction of potentially greater use of administrative records for medical care statistics. For person level analyses, it assumes that accurate measures are possible of the number of different medical providers with whom each sample individual in the population of interest has had contact; that is, a measure of multiplicity which is essential to appropriate weighting of the individuals ultimately selected for the sample from the medical provider files. As a design, it has other deficiencies since it does not produce a complete medical care history for the period of interest for the sample individuals. Thus, for example, distributions of individuals by number of medical provider visits in a specific time period would not be possible. Estimates of the total number of visits made to medical providers, the services provided and the charges can all be generated in total and separately for population subgroups based on the characteristics of the sample of persons making the visits. Person level analyses should also be possible, using the multiplicity adjusted weights, but the variance of person level estimates can be expected to be considerably larger than for self-weighted samples of equivalent size. Also it should be noted that population based statistics, such as rates, are still not possible without statistics on the characteristics of the population at risk.

Still, this type of design deserves in-depth study. The pressure for greater use of administrative records in survey research can be expected to increase in the future as data collection costs for population based samples increase. The survey design challenges implicit in such pressure should be addressed now, in my view.

VI. Estimation

The statistician faced with combining household survey data with record check data to produce more accurate estimates than either data set can provide separately has a number of choices. The alternatives deserve more study and comparison, particularly dual system, imputation and weight adjustment procedures. The assumptions implicit in each procedure need to be listed and compared together with specification and assessment of the magnitude of each of the different components of error generated by each procedure. Finally, appropriate methods for estimating the mean square error associated with estimates based on a specific procedure for combining the two (or more) data sets need to be developed. These methods must identify the potential sources of error in each procedure, for example, the added variance due to imputation, and assess their relative magnitudes.

REFERENCES

1. Andersen, R., Kasper, J., Frankel, M. R. and Associates. Total Survey Error. San Francisco, Calif. Jossey-Bass, Inc., 1979.
2. Folsom, Ralph E., Rick L. Williams, and James R. Chromy (1980). Optimum Design of a Medical Care Expenditure and Utilization Survey Involving a Provider Record Check: An Application of Survey Cost Minimization Subject to Multiple Variance Constraints. RTI Report No. RTI/1725/01-065.
3. Marquis, Kent H., Record Check Validity of Survey Responses: A Reassessment of Bias in Reports of Hospitalizations. R-2319-HEW, Rand Corporation, Santa Monica, Calif., May 1978.