

ITEM NONRESPONSE OF MEDICAL PROVIDER UTILIZATION DATA IN THE NMES INSTITUTIONAL SURVEY

D.E.B Potter, Jill J. Braden, Agency for Health Care Policy and Research
D.E.B. Potter, AHCPR, 2101 E. Jefferson Street, Suite 500, Rockville, MD 20852

KEY WORDS: Nonrespondent characteristics, questionnaire design, nursing homes, facilities for persons with mental retardation

The Institutional Population Component of the 1987 National Medical Expenditure Survey (NMES) was established to provide an assessment of the utilization, costs, sources of payment and health status of the U.S. population living in nursing and personal care homes and facilities for persons with mental retardation. The period of assessment covered calendar year 1987, during which time data were collected for a sample of residents and admissions to nursing and personal care homes and facilities for persons with mentally retardation.

This paper adds to the previous research investigating nonresponse in NMES Institutional Population Component (IPC) by evaluating the item nonresponse associated with the utilization of medical providers (e.g., physicians) by persons institutionalized in a nursing or personal care home or a facility for persons with mental retardation, an area not yet examined for the NMES IPC. The results of this analysis will be used in the development of an estimation strategy for use with the NMES IPC medical provider data and to inform the design of the next NMES Institutional Population Component, currently planned for 1996.

Data Sources

The NMES Institutional Population Component was a year-long two-stage panel survey (Cohen, Potter and Flyer, in press). In the first stage, facilities - nursing and personal care homes (NHs) and facilities for persons with mental retardation (FMRs) - were selected. Residents on January 1, 1987 and admissions (persons admitted between January 1, 1987 and December 31, 1987) were sampled in the second stage. The combination of the sample of residents on January 1 and the admission sample, represents the 1987 population

in nursing or personal care homes or in facilities for persons with mental retardation.

To meet the strict estimation requirements imposed on all NMES surveys, the IPC sample used for this analysis was restricted to sampled persons (SPs) who had expenditure and facility use data for at least a third of their period of 1987 eligibility (Cohen, Potter and Flyer, in press). Sampled SPs with a prior 1987 stay in a non-sampled eligible place have also been excluded from this analysis. As a result, this analysis is based on 8,992 persons (5,071 in NHs and 3,920 in FMRs), representing 2,242,210 persons in nursing and personal care homes and 224,817 persons in facilities for the mentally retarded.

The operational implications of selection of both January 1 residents and admissions over the course of 1987 required several trips to each facility (Edwards and Edwards, 1989). Interviewers made four visits to each co-operating facility at approximately four month intervals to facilitate sample selection and data collection.

Data collected from facility respondents included facility characteristics, physical and mental health status of sample persons, and socio-demographic characteristics, residence history information, institutional use and expenditures, and facility reported medical provider use. The data on medical provider use (and facility use and expenditures) were collected with the Institutional Use and Expenditure Questionnaire (IUEQ). The IUEQ was administered to facility staff during the last three rounds of IPC data collection. Because of the requirement that responding SPs have at a minimum at least a third of their use and expenditure data, all responding SPs have at a minimum, their first IUEQ. In this first analyses of medical provider item nonresponse, data collected with an SPs first IUEQ are examined.

In total, data were collected on use for 14 different types of medical providers. For each provider type, the facility respondent was asked if the SP had received care from the provider during their period of institutionalization (between the

start and the end date of the round). The respondent was instructed to include utilization that the SP received outside of the facility as well as utilization from any providers seen inside the facility (see Edwards and Edwards, 1989). If care was known to have been provided to the sampled person, the respondent was further asked how often was the care provided. If the number of provider visits was known and the provider was not a hospital-based provider, the respondent was further asked how many of those visits were paid for by the facility. The 14 medical provider types are shown below in the order of their appearance in the questionnaire:

- Dentist, dental surgeon, orthodontist, dental assistant, or other dental professional;
- Psychiatrist, psychologist, psychiatric social worker, psychiatric nurse, or other trained professional for mental health care;
- Hospital emergency room;
- Hospital outpatient department or hospital clinic;
- Medical doctor (of any kind);
- Optometrist;
- Podiatrist;
- Chiropractor;
- Physical therapist;
- Occupational therapist;
- Speech or hearing therapist;
- Respiratory therapist;
- I.V. Therapist;
- General or short term hospital.

Item Nonresponse Rates

Item nonresponse for these 14 types of medical providers was evaluated. Rates of nonresponse were found to be very low for hospital-based provider items. Rates were higher for items related to non-hospital-based providers, but still relatively low (with a mean nonresponse rate of 1 percent). However, nonresponse rates were more pronounced for the follow-up questions. For example, for the follow-up questions on amount of medical provider utilization, nonresponse was more than triple, on average, that found for the initial utilization items. Cumulative nonresponse rates (across the initial utilization item and the two follow-up questions) were even higher and averaged eight times that of the initial utilization

items (with a mean of 8.5 percent).

It was hypothesized that the rate of nonresponse might vary across subdomains by provider and item type. The sheer numbers of medical provider items (potentially as many as 84 questions) prohibited an in-depth analysis of each item. Consequently, analyses were limited to two nonhospital-based provider types: physicians and physical therapists. The physician items were selected, in spite of the low rates of nonresponse, because of the importance physicians have in the delivery of health care. The physical therapist (PT) items were chosen because of their higher rates of nonresponse in comparison to physicians and because, when compared to the other medical provider items with similar rates of nonresponse, the sample sizes for the PT items were larger. Five items were examined in detail; initial physician utilization, amount of physician utilization, and number of physician visits paid by the facility, the initial utilization item and amount of utilization item for physical therapists (PT).

It was hypothesized that each of the three item types (i.e., 1-any utilization, 2-amount of utilization, 3-amount of facility paid utilization) would be associated with different correlates of nonresponse. In total, 52 potential correlates of nonresponse were examined, for each of the five medical provider items. The correlates examined fell into seven distinct groups. These were:

Socio-demographic characteristics of the sampled person (SP), e.g., age, race, living kin, and income.

Characteristics of the SP's facility charge: sources of payment for facility charges at the time of data collection and a measure for the SP's per diem rate charge.

Residence history characteristics such as whether the SP was in the facility at the time of data collection, how long the SP had resided in the facility at the time of data collection and the SP's location prior to admission.

Health status measures, e.g., did the person have ADL difficulty (bathing, feeding, etc.), number of behavioral problems and number of psychiatric symptoms.

Facility characteristics, e.g., facility size, services and location; and ownership type.

Interviewing environment, e.g., the facility respondent's job title, length of the reference period and number of SPs requiring data collection in the facility.

Interviewer characteristics such as age, race, education, and years of survey experience.

The descriptive analysis involved running Chi-square statistics for each of the five medical provider items (each classified as responding versus nonresponding) against each of the 52 potential correlates. The chi-square test of homogeneity was used to assess significance. The statistics were run weighted using the RTI SUDAAN software (Shah et al, 1992), which normalizes the sampling weights and adjusts for the complex survey design of the NMES institutional survey. These analyses were run weighted due to the disproportionate sampling in the NMES IPC as a result of the multi-stage design, nonresponse and post-stratification adjustments, and because the intent was to generalize the results to other national surveys of the institutionalized. All variables with a p value of .20 or less for a medical provider item were included in the initial model specification for that item, as described below.

Models to Predict Medical Provider Item Response

Weighted logistic regression analyses were run for four of the five medical provider items. A separate model was developed for each item. Results for three models; Any physician utilization, Number of visits to a physician and Any physical therapist utilization; are presented here. The initial model specifications for the independent variables in each of the three models were: facility type (NH or FMR), sample type (January 1 resident or New admission), and all other variables in the chi-square analysis with a p value of .20 or below.

To identify variables requiring exclusion from the model due to multicollinearity of the independent variables, correlation analysis were

performed. Based upon these results, the variables on number of ADLs, and sample type (found to be correlated with the individual ADL variables and length of stay in the facility, respectively) were eliminated from all model specifications. Also eliminated as independent variables, at this time, were all interviewer variables. With these specifications, the initial models to predict item response were run as backwards stepwise logistic regression models.

The results of the initial modeling runs were used to specify a second set of logistic models, also developed to predict item response. The specifications for the second set of models were: (1) the "final" models as determined from the previous backwards logistic regressions, (2) the inclusion of interviewer variables and (3), the exclusion from the analysis of all persons (289) whose medical provider data were collected by a supervisor, rather than an interviewer. The exclusion of the supervisor cases served several purposes: (1) it eliminated the multicollinearity problems encountered initially with the interviewer variables, (2) it controlled for the difference in item response rates between supervisors and interviewers while (3), simultaneously controlling for the fact the supervisors might have more difficult cases than interviewers (the transfer of a case from interviewer to supervisor typically occurs as part of the refusal conversion process). The results of the last set of specifications for weighted backwards logistic regression models provided the final models to predict response. These models were rerun to adjust for survey design complexities (Shah, et al 1992). The resulting models are shown in Table 1.

There were some consistent findings across the three models. As a group of variables, the socio-demographic characteristics of the sampled persons were not that important in predicting item response. There was a finding for one of the models (the initial PT utilization model) that when race of the SP was other than white, there was an association with a decrease in response. A higher level of income (>\$10 to \$20K) was found to be associated with an increase in response (in the physician visit model). And in the same model, persons with the lowest per diem rates (less than \$35/day) for facility charges were found to be associated with lower item response profiles.

However, persons in the next rate category (\$35 - \$50/day) were found to be associated with an increase in the probability of response (in the initial PT utilization model).

Two measures of residence history were found to be predictors of response in two of the models (the model for number of physician visits and the PT utilization model). Persons with a long length of stay in the facility at the time of data collection (stays greater than 3 years) were found to be associated with higher probabilities of response. Similarly, persons not in the facility at the time of data collection were found to be associated with lower probabilities of response.

Health status measures were found not to be important predictors of medical provider item response in the three models evaluated. There was a single significant finding, in both of the initial utilization models, that persons with no mental disorders (e.g., no senile dementia) were associated with a lower probability of response.

Some of the facility characteristic measures were found to be predictors of response, but there was little consistency across models with the exception of a variable that measured the facility's provision of physical therapy services by trained licensed professionals. In all three models, item response was increased for SPs in facilities providing these services. Facility location in a MSA (in the physician utilization model), and the facility's provision of physician services at least 40 hours a week (in the physician visits model) were associated with a decreased response profile, while small facility size (in the physician visit model) was found associated with increased response. Across all models, facility type (NH or FMR), was consistently found not to be a predictor of item response.

In all three models, it was clear that the probability of response was associated with the job title of the respondent who provided the medical provider data and the number of respondents to the questionnaire containing the medical provider data. Data provided by persons providing direct patient care (e.g., nurses) or their immediate supervisors (e.g., Directors of Nursing) were associated with higher probabilities of response. Conversely, respondents in a residual "other" job title category (i.e., those who were not providers of medical or social services, facility administrators, or medical

records personnel) were less likely to be associated with response.

In both of the initial utilization models, sampled persons whose questionnaire data were collected from two respondents had a greater association with response. Other interviewing environment variables found to be significant predictors of response were measures for the length of the reference period; however, these results were inconsistent across models.

Interviewer tenure was a significant predictor of response in the three models. Fewer years of interviewing experience was consistently associated with a decrease in the probability of response. The age of the interviewer (≤ 39), was an important predictor of response (in the physician utilization model); and the coefficient was more than four times that of many of the other variables in the model. Race of the interviewer other than white was found to be associated with a significant decrease in the probability of response, and similar to the relationship of age on prediction on the same outcome measure (i.e., physician utilization), the coefficient was relatively large in comparison to the other variables in the model. Interviewers education was also a predictor of response, but the results were more difficult to interpret; a little college education was associated with an increase probability of response, while a post college education was found to be associated with a decrease in response.

Discussion

The modeling results suggest the following:

- (1) In spite of the low level of nonresponse for the initial physician utilization item, the error is not purely random.
- (2) Several predictors of item response were consistently identified across multiple models. Found to be significant predictors of medical provider response were: interviewer and respondent characteristics, the sampled person's residence history, the SP's lack of mental disorders, and the facility's provision of PT services by trained licensed professionals.

(3) While there was consistency of results across models, each model contained some unique significant predictors not found in other models. This suggests that the mechanisms used by facility respondents to recall medical provider data, whether recall is from memory or records, differ depending upon the types of utilization data and the type of provider.

No recommendation is made at this time for major changes to the data collection methods used to collect these medical provider data. The rates of nonresponse are relatively low for the initial utilization variables, and the contribution to bias caused by this nonresponse should be small. However, these results do offer some insight into issues of data quality in institutional surveys that are in contrast to household surveys or previous findings from research on the NMES institutional survey.

The type of facility, NH or FMR, was found not to be associated with medical provider item nonresponse. This result is in contrast to results from previous NMES IPC research which found unit response (response to the entire questionnaire, rather than items in the questionnaire) to be associated with facility type (Tourangeau and Johnson, 1992).

The only health status parameter found to be significant across models was a variable that measured the absence of mental disorders. Perhaps there is increased saliency for respondents to recall (whether from memory or records) medical provider utilization data when the sampled persons lack the ability to remember utilization for themselves (i.e., because the SP has dementia). Thus, facility personnel (respondents) cannot rely on the SP for this information, and develop techniques to store (in either memory or records) this information.

These results provide evidence that the probability of item response in institutional surveys may be associated with the saliency of the sampled person (not the salience of the event) in the mind of the respondent. Persons no longer in the facility were associated with a lower probability of response, while persons with long stays were associated with higher levels of response.

Burden (measured as the number of sampled persons requiring data collection) was not found to

be associated with item response in this analysis. In addition, the use of multiple respondents (a measure thought to increase burden) was associated with an increment in the probability of response. Two possible explanations are offered for these findings. (1) As the data collection task becomes more routinized (i.e., more SPs requiring data collection), the less difficult the task becomes for the respondents (or the interviewers). Thus, an increase in number of SPs would not be associated with a decrease in the probability of response. (2) These analyses are conditioned on the facility agreeing to participate in the survey. Perhaps, once a facility agrees to participate, survey burden is no longer an important issue -- it is only an issue for gaining initial facility cooperation i.e., for unit nonresponse not item nonresponse.

These results provide some evidence that response to medical provider utilization items is based upon knowledge that respondents recall and that this knowledge is not enhanced by the use of medical records. But, the use of medical records to retrieve information on the amount of utilization (for physician visits) is associated with an increase in item response.

It is clear from the analyses that locating the "best" respondent within the institution has the potential to significantly increase rates of response. Clearly, interviewer training for surveys of the institutionalized needs to teach interviewers how to "navigate" within institutions to find the most appropriate respondent.

Lastly, it was not possible to evaluate the level of quality for the data reported; high rates of response do not guarantee high data quality. Of particular interest is the accuracy of facility reporting on the number of physician visits. Institutionalized persons can, and do, receive some level of physician care outside of the institution. Are the total number of physician visits reported by facility respondents under reports of use? It might be proposed to examine the reporting of physician utilization for the all-year Medicare Beneficiaries in the NMES IPC compared to that reported in administrative records on Medicare Part B claims data. Unfortunately, the claims data capture only the cost of utilization and not the amount of utilization. Future research to evaluate the quality of institutional physician utilization data will first involve overcoming this limitation.

Table 1. Coefficients for logistic regression models to predict medical provider item response

Independent variables	Model coefficients		
	Physician models		Physical therapy model
	Any utilization	Total number of visits) of those with utilization)	Any utilization
Intercept	5.8432	2.5123	5.8274
<u>Socio-Demographic Measures</u>			
Race of SP is not white	-0.9864**
No living siblings	0.7196
Income \$10 - \$20K	...	1.3730*	...
<u>Characteristics of Facility Charges</u>			
Per diem charge < \$35/day	...	-0.6790*	...
Per diem charge \$35-\$50/day	1.0380*
<u>Residence History Characteristics</u>			
Length of stay in facility ≥3 years	...	0.6910*	0.6667*
Non-resident during data collection	...	-1.0566**	-0.7620**
<u>Health Status Measures</u>			
Has difficulty dressing	0.7422
Has no dementia/mental disorders	-0.8698**	-0.5328	-0.6962**
<u>Facility Characteristics</u>			
Small facility size (NH: <50 beds; FMR: <16 beds)	...	2.2974**	...
Physician on-site ≥40 hr/wk	...	-1.3799**	...
Located in an MSA	-1.6426*	...	-0.9187
PT provided by licensed professional	0.9778*	1.5681**	1.3514**
<u>Interviewing Environment</u>			
Job title of respondent:			
Director Nursing/Health Services	...	1.8789**	...
RN, LPN or other care giver	...	1.2451**	...
Medical records personnel	...	0.9486	...
Other (but not Administrators)	-2.2874**	...	-1.7140**
Quest w/ med provider data administered to 2 respondents	0.8459*	...	1.0217*
Medical records used during interview	...	1.0226**	...
Length of reference pd 91-180 days	-0.6270
Length of reference pd 181-365 days	0.9236*
<u>Interviewer Characteristics</u>			
I's age 20 - 39 years	4.5441*	...	1.6325
I's race is not white	-3.1880*	...	-1.9068
I has some college	0.9546	0.8766**	...
I has post BS education	-0.7696
I has 0 years survey experience	...	-1.0374	...
I has 1-4 years survey exp	-1.8757**	...	-1.3758**
I has 5-9 years survey exp	...	-1.2885**	-0.8129
I has no experience w/ institutional surveys	-0.9542
I's exp w/ institutional surveys is unknown	-1.603*	...	-0.866
Model Chi-square and p value	263 .0001	199 .0001	269 .0001
Unweighted N	8,703	7,733	8,703

*p <.05, **p <.01, "... indicates that the variable not included in final model. Source: Agency for Health Care Policy and Research. 1987 National Medical Expenditure Survey - Institutional Population Component.

The authors thank Steven B. Cohen, AHCP, for his insightful comments, and Jane Faulman and Chao-Sung Yu for invaluable computer processing support. Readers who wish to obtain a comprehensive version of this paper and/or additional tables and references should contact the author. The views expressed in this paper are those of the authors and no official endorsement by the Department of Health and Human Services or the Agency for Health Care Policy and Research is intended or should be inferred.