PATTERNS OF CONDITIONAL NONRESPONSE IN THE MEPS PARENT ADMINISTERED QUESTIONNAIRE (PAQ)

Frances M. Chevarley, Agency for Healthcare Research and Quality
Frances M. Chevarley, AHRQ, 2101 E. Jefferson Street Suite 500W, Rockville, MD 20852

Key Words: MEPS, Parent Administered Questionnaire, PAQ, conditional nonresponse rate

1. Introduction

The Parent Administered Questionnaire or PAQ was given out for each child in the family in advance of the fielding of round 2 of panel 5 and round 4 of panel 6 of the Medical Expenditure Panel Survey (MEPS) during the second half of the year 2000. The PAQ along with a corresponding Self Administered Questionnaire (SAQ) for each adult in the household were designed as enhancements to MEPS to measure health care quality for children and adults, respectively. In this paper we explore patterns of unit nonresponse for the PAQ in terms of demographic and other variables in the core MEPS questionnaire conditional on response to rounds one and three in 2000 (that is conditional on those cases in the 2000 Point-in-Time file). We are interested in knowing what types of children are more likely to have their questionnaires completed or not. Besides descriptive analysis, logistic regression was also used in order to determine the primary factors that differentiate those children with completed PAQ's versus those whose PAQ was not completed. This analysis was the first step in producing nonresponse adjusted weights for this PAQ.

2. Background

2.1 Medical Expenditure Panel Survey (MEPS)[1]

The Medical Expenditure Panel Survey (MEPS) was designed to provide nationally representative annual estimates of health care use and expenditures, access to care, patient and customer satisfaction, health status, and insurance coverage for the U.S. civilian noninstitutionalized population. It is co-sponsored by the AHRQ and co-sponsorship from the National Center for

Health Statistics (NCHS).

MEPS is the third in a series of national probability surveys conducted by AHRQ on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES) was conducted in 1977, and the National Medical Expenditure Sruvey (NMES) in 1987. Beginning in 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance system.

The MEPS actually comprises a family of surveys that cover three different components of the U.S. health care system: the Household Component (HC), the Medical Provider Component (MPC), and the Insurance Component (IC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications. This paper focusses on the PAQ which is a special supplement to the MEPS-HC.

2.2 MEPS Household Component (HC)

The MEPS-HC is a nationally representative survey of the U.S. civilian noninstitutionalized population. It uses the National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics as its sampling frame. The NHIS sampling frame provides a nationally representative sample of the U.S. civilian non-institutionalized population and reflects an over sampling of Blacks and Hispanics.

The HC collects medical expenditure data at both the person and household levels. The HC uses an overlapping panel design in which, for each panel, data are collected through a preliminary contact followed by a series of five in-person interviews over the course of a two-and-a-half year period of time (2 interviews per year). Using computer assisted personal interviewing (CAPI), in-house personal interviews are conducted by Westat Interviewers. The HC survey is long and each of the five interviews averages about 1 ½ hours in length. The MEPS HC interview typically is conducted with a household respondent, who answers for him/herself and other household members. Because certain information can only be obtained reliably through self rather than proxy-response, several self-administered questionnaires are also used over the 2 ½ year period to collect information for

each household member. Parents or guardians are asked to complete parent administered questionnaires for their children (aged < 18).

The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

2.3 MEPS-HC Enhancements to Measure Healthcare Quality

Because of AHRQ's mandate to Measure Healthcare Quality for a National Quality Report starting in 2003, AHRQ implemented a number of enhancements to measure healthcare quality beginning with the 2000 MEPS-HC. The 2000 and 2001 enhancements to MEPS included three paper questionnaires: the Parent Administered Questionnaire (PAQ) which was to be completed by the parent for each child in the family; the Self-Administered Questionnaire which was to be completed by each adult in the family; and a Diabetes Self-Administered Questionnaire which was to be completed by each person (or parent for children) in the family with Diabetes. Additional enhancements included adding questions for quality priority conditions and also a preventive care supplement to the 2001 MEPS-HC questionnaire. Although the quality priority conditions, the preventive care supplement, and the Diabetes Self Administered Questionnaire were collected in 2001 (rounds three and five of panels five and six, respectively) they relate to the same respondents and estimates in 2000 (earlier rounds of panels five and six, respectively) and will be analyzed with the other enhancements from 2000.

2.4 MEPS-HC Parent Administered Questionnaiare (PAQ) in 2000

The 2000 MEPS-HC Parent Administered Questionnaire (PAQ) was a very brief paper questionnaire that contained thirteen questions from the CAHPS, the Consumer Assessment of Health Plans Survey--a survey tool that measures health care quality information from a consumer perspective. The 2000 PAQ also contained 14 Living with Illness Measure (LWIM) questions that enabled the differentiation of quality care and costs of children with and without chronic illnesses and disabilities.

Starting in 2001, the questions from the 2000 PAQ were merged into the MEPS-HC questionnaire in a new Child Health Supplement (CHS) that would be administered in rounds 2 and 4 (the second half of a typical year) beginning in 2001. In addition to CAHPS and LWIM questions, the CHS includes questions from the Resistence to Illness measure and the Columbia Impairment Scale.

3. Methods

3.1 MEPS Overlapping Panel Design

The MEPS-HC collects data through an overlapping panel design. In this design, two calendar years of information are collected from each household through five in-person interviews, two each year (Figure 1). This series of data collection activities is repeated each year on a new sample of households resulting in overlapping panels of survey data. As can be seen in Figure 1, during the first half of a typical year, there are three rounds of data in the field at the same time. For the first half of 2001, round five of the 1999 panel, round three of the 2000 panel, and round one of the 2001 panel were in the field. During the second half of 2001, and of any typical year, two rounds of data are in the field at the same time. During the second half of 2001, round four of the 2000 panel and round two of the 2001 panel were in the field at the same time.

Figure 1
MEPS Panel Design: Data Field Periods
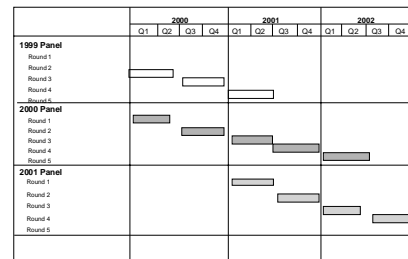
3.2 MEPA Public Release Files (PUFs)

Data release products from MEPS-HC were designed to combine subsequent panels in order to produce more precise estimates for monitoring health care utilization and expenditures. Public Use Files (or PUFs) for the MEPS-HC are generally of two types: the point-in-time file (PIT file); or Full Year Files (FY files). The point in time file is produced after the data from the first half of the year are processed and is based on data from round one of the one year's panel and round three from the previous year's panel for events in the first half of the year. Round five, although also collected during the first half of the year, is not included because the reference period for round five is the second half of the previous year and therefore cannot be used to make estimates for the first half of the current year. The Point in Time (PIT) file then is based on a half of a year's worth of data from rounds one and three, has quick turnaround and does not contain the full set of variables.

On the other hand, Full-Year files are based on a full year's worth of data collection from rounds one, two, and three of one panel; and rounds three, four and five of the previous year's panel. The Full Year file takes longer to produce, and contains many more variables.

3.3 PAQ file matched to the 2000 PIT

During 2000, in advance of the round two interview of the 2000 panel and the round four interview of the 1999 panel, PAQs for each child in the family were mailed to the household for the parents to complete. The completed PAQ's would then either be picked up by the interviewer during the rounds two and four interviews or would be mailed back by the respondent. The PAQ questionnaires that were not received by the end of the rounds two and four interviews , were prompted for again in CAPI at rounds three and five (in the first half of 2001). If a person said they lost or misplaced their PAQ when interviewers where checking on the status of the PAQs (rounds two and four or rounds three and five), then the interviewer would give them another and hand write the label information.

Because the PAQ information are needed for the 2003 National Quality Report, the PAQ information was matched back to the socio-demographic and other information of the 2000 Point-in-Time file, which is processed much more more quickly then the 2000 Full Year file. At a later time when the 2000 Full Year file has been processed, the PAQ information will be merged with the 2000 Full Year file to produce a Public Use File.

Again because of the need to analyze these data as quickly as possible, a preliminary PAQ file and not the final PAQ file was used for the analysis in this paper because the final PAQ file was not available when this analysis was done. This preliminary file contained about 85-90 % of the coded PAQs.

The published 2000 Point-in-Time file contained 7,363 records of children (ages < 18) with positive person weight. Of these, 5 children became out-of-scope for receiving the PAQ during rounds one and three; they either died, moved out of the country, or were institutionalized. Our best estimate of the number of children eligible to receive the PAQ questionnaire is based on the 7,363-5=7,358 children in rounds one and three in 2000.

Now the PAQ preliminary file had 6,978 completed PAQs, of which 231 did not match back to rounds one and three (the 2000 PIT), 292 cases had a weight of zero in the 2000 PIT, and one case had age = 18, which would make it out-of-scope for the PAQ. This resulted in a preliminary file of 6,978-231-292-1=6,454 PAQ records that matched back to rounds one and three (the 2000 PIT) and who had positive weight in the 2000 PIT.

3.4 Data analysis

The goal of this analysis is to examine attrition in the PAQ conditional on response to rounds one and three in 2000 (the 2000 PIT file); provide a profile of respondents and nonrespondents to the PAQ for nonresponse adjustment to the weights; and give feedback to the interviewers on characterisitics of non-respondents to the PAQ.

To examine attrition in the PAQ and to provide a profile of responders and nonresponders to the PAQ for nonresponse adjustment to the weights, the preliminary PAQ file was merged with the MEPS 2000 PIT file in order to obtain socio-demographic characteristics of the PAQ responders and non-responders.

It's always the proxy that answers for the child, but the nonresponse estimates are targeted to children's socio-demographic characteristics because we need to know the end result of that interaction to make non-response adjusted weights. The child's socio-demographic variables that were analyzed with respect to PAQ nonresponse were: region of the country; whether Metropolitan Statistical Area or not; family size; gender; age; race/ethnicity; health status; and health insurance.

The general household characteristics are based on the reference person's variables (the person who owns or rents the residence). The reference person's characteristics give a flavor about the socio-demographic aspects of the family. The reference person's characteristics that were analyzed were: age, gender, race/ethnicity, whether a spouse was in the house; education; health status; mental health status; employment status; and health insurance. Future efforts will include analyzing response rates for the respondent to help with informing the field effort, but our efforts here focus on the characteristics of the child and of the reference person.

As a first step, bivariate nonresponse was examined for each of the child's and reference person's characteristics mentioned above using SUDAAN's Proc Crosstab. Those variables with a Chi-square value $\leq .05$ were considered to be associated with PAQ nonresponse.

In order to ascertain primary factors that differentiated the PAQ responders from the nonresponders, weighted logistic regression analyses were conducted using SUDAAN's Proc Logistic. All of the variables that were considered for the bivariate analysis that compared the profiles of responding and non-responding PAQs were initially considered. A backward elimination technique was implemented in order to identify the core set of factors that were determined to be significant predictors of survey response status at the alpha=.05 level in the logistic regression model under consideration.

SUDDAN was used for both the bivariate and multivariate analysis in order to account for the MEPS survey design complexities. The 2000 PIT file weights that are unadjusted for the PAQ nonresponse were used in both the bivariate and multivariate analysis.

4. Results

The estimated response rate conditional on the existence of a child in the 2000 PIT file for this preliminary PAQ file was 6,454/7,358=87.7%. This estimated response rate is lower than the actual PAQ response rate because the preliminary PAQ file was missing some of the coded PAQs. It is also low because

the number of eligibles from rounds one and three may be an overestimate of eligibles because additional children may become out-of-scope during rounds two and four (the rounds in which the PAQ was administered). The response rate for the PAQ estimated from the field was in the range of 95-96%.

The bivariate nonresponse analysis was examined for each of the child's and reference person's characteristics previously listed.. After analyzing nonresponse for each of these variables, only three were found to have a significant Chi-square value. The bivariate analysis results for the child variables are shown in Table 1 at the end of this paper.

As shown in Table 1 and Figure 2, children living in MSAs were more likely to have higher PAQ non-response. The PAQ was also less likely to be completed for older children, ages 12-17, then for children of the other age groups. The PAQ was also less likely to be completed for the child when the reference person was 60 years of age or older than when the reference person was less than 30 years of age.

In order to ascertain primary factors that differentiated the PAQ responders from the nonresponders, weighted logistic regressions were conducted using SUDAAN. All of the variables that were considered for the bivariate analysis that compared the profiles of responding and non-responding PAQs were initially considered and a backward elimination technique was implemented. Those variables associated with the child that were found to be significant

Figure 2
PAQ Percent Non-response
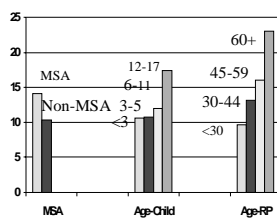
MEPS 2000 PIT_PAQ

in the final model are shown in Figure 3. The factors associated with the child that were in the final model were MSA status , child's age, and child's health status. The '*' indicates that the estimated odds ratios were found to be significant at the alpha=.05 level. As shown in Figure 3, children living in MSAs, with an estimated odds ratio of 1.44 were 1.44 times more likely not to complete the PAQ than children not living in MSAs. Younger children were less likely not to complete the PAQ than older children (aged 12-17), and children with very good reported health status were less likely than children of good, fair and poor health status combined not to complete the PAQ.

5. Conclusion

Based on our analysis of a preliminary PAQ file conditional on response to the MEPS rounds one and three interview, higher nonresponse was found in Metropolitan Statistical Areas and for older children. Lower nonresponse was also found for children of very good reported health status versus good,, fair, or poor combined.

Figure 3
Logistic Regression Results

■ MSA			■ Child's Health Status	
– MSA	1.44*		– Excellent	0.81
– non-MSA	1.00		– Very Good	0.66*
■ Child's Age			–Good,Fair,Poor	1.00
– <3	0.56*			
– 3-5	0.56*			
– 6-11	0.64*			
– 12-17	1.00			

Based on these analyses nonrespone adjusted weights for the merged PAQ file with the 2000 PIT file have subsequently been produced. We are therefore able to move ahead with the editing of the PAQ and other enhancement variables and with the table production for the National Quality Report.

REFERENCE

[1] Joel W. Cohen, Ph.D., Alan C. Monheit, Ph.D., Steven B. Cohen, Ph.D, and Doris C. Lefkowitz, Ph.D., et al, *The Medical Expenditure Panel survey: A National Health Information Resource,* AHCPR Pub. No. 97-RO43, March 1997

Table 1
Estimated number and percent distribution with and without completed PAQ and PAQ nonresponse for Children (under 18 years of age) according to selected child variables: 2000 MEPS

Variable	Number of children (in thousands)	Percent distribution (95% Confidence Intervals)		Percent PAQ Nonresponse (95% Confidence Intervals)
		With Completed PAQ	Without completed PAQ	
Region		100%	100%	
Northeast	13,127	17.6 (15.0-20.5)	22.3 (17.2-28.3)	16.3 (12.6-21.0)
Midwest	17,462	24.6 (21.5-28.0)	21.8 (17.4-27.0)	12.0 (9.8-14.7)
South	24,117	34.1 (30.0-38.3)	29.8 (24.6-35.6)	11.9 (9.6-14.6)
West	17,308	23.7 (19.0-29.2)	26.1 (20.0-33.4)	14.5 (12.1-17.4)
		Chi-Square = 4.93 (3df), p=.3667		
MSA		100%	100%	
MSA	58,511	80.6(76.7-84.0)	85.5 (80.6-89.4)	14.1 (12.4-15.9)
non-MSA	13,502	19.4 (16.0-23.3)	14.5 (10.6-19.4)	10.3 (8.0-13.3)
		Chi-Square = 4.86 (1df), p=.0287		
Family Size		100%	100%	
1- 3 persons	18,442	25.8 (23.9-27.8)	24.3 (20.4-28.6)	12.7 (10.6-15.1)
4-person	24,293	33.7 (31.8-35.5)	34.3 (29.5-39.4)	13.6 (11.7-15.7
5+ persons	29,278	40.5 (38.0-43.2)	41.5 (35.9-47.2)	13.6 (11.2-16.6
		Chi-Square = 0.51 (3df), p=.9161		
Gender		100%	100%	
Male	37,135	51.3 (49.9, 52.8)	53.1 (49.4, 56.6)	13.8 (12.4-15.3)
Female	34,878	48.7 (47.2, 50.1)	46.9 (43.4, 50.6)	13.0 (11.2-15.0)
		Chi-Square = 0.82 (1df), p =.3667		

Table 1
Estimated number and percent distribution with and without completed PAQ and PAQ nonresponse for Children (under 18 years of age) according to selected child variables: 2000 MEPS

Variable	Number of children (in thousands)	Percent distribution (95% Confidence Intervals)		Percent PAQ Nonresponse (95% Confidence Intervals)
		With Completed PAQ	Without completed PAQ	
Age		100%	100%	
0-2	11,380	16.3 (15.1, 17.5)	12.7 (9.8, 16.4)	10.7 (8.3-13.8)
3-5	12,637	18.1 (16.8, 19.4)	14.2 (11.8, 17.1)	10.8 (8.8-13.3)
6-11	24,439	34.5 (33.0, 36.0)	30.4 (27.0, 34.1)	12.0 (10.0-14.3)
12-17	23,557	31.2 (29.5, 32.9)	42.7 (38.2, 47.3)	17.4 (15.3-19.3)
		Chi-Square = 22.40 (3df), p=.0001		
Race/Ethnicity		100%	100%	
White, non-Hispanic	49,218	68.2 (65.0, 71.2)	69.4 (63.8, 74.6)	13.6 (11.9-15.5)
Black, non-Hispanic	11,104	15.2 (12.6, 18.3)	16.5 (12.3, 21.9)	14.3 (11.0-18.4)
Hispanic	11,691	16.6 (14.0, 19.5)	14.0 (11.1, 17.6)	11.6 (9.6-13.9)
		Chi-Square = 2.43 (2df), p=.2988		
Health Status		100%	100%	
Excellent	37,593	52.0 (50.1-53.9)	53.5 (48.8-58.0)	13.7 (12.0-15.6)
Very Good	2,031	29.7 (28.1-31.4)	26.0 (22.2-30.1)	11.9 (9.8-14.4)
Good/Fair/Poor	13,385	18.3 (16.6-20.1)	20.5 (16.9-24.7)	14.8 (12.2-17.8)
		Chi-Square = 2.82 (3df), p=.4220		
Health Insurance Coverage				
Yes	62,011	86.3 (84.3-88.00	85.1 (81.1-88.4)	13.2 (11.7- 14.90
No	10,002	13.7 (12.0-15.7)	14.9 (11.6-18.9)	14.4 (11.7-17.5)
		Chi-Square = 0.45 (1df), p=.5055		

SOURCE: Author's calculations based on SUDAAN's Proc Crosstab using the Medical Expenditures Panel Survey, preliminary 2000 PAQ file merged with the 2000 Point-in-Time file.