The Utility of the Integrated Design of the Medical Expenditure Panel Survey (MEPS) to Inform Trends in Nonresponse¹

Frances M. Chevarley, Ph.D. Karen E. Davis, M.A.

Agency for Healthcare Research and Quality Department of Health and Human Services 540 Gaither Road, Rockville, MD 20850

Proceedings of the 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference

Abstract

This paper analyzes estimates from the Medical Expenditure Panel Survey (MEPS) matched with the National Health Interview Survey (NHIS) and uses practical tools to inform MEPS nonresponse estimates. MEPS is a nationally representative panel survey studying health care use, access, expenditures, source of payment, insurance coverage, and quality of care. Each year a new panel begins and each panel has 5 rounds of data collection over 2 ½ years that cover a two-year period.

The goal of this paper is to inform trends in MEPS nonresponse. Because MEPS uses the NHIS, conducted by the National Center for Health Statistics as its sampling frame, estimates are produced using variables from both the NHIS and MEPS across different categories of the MEPS response categories. Data used are from the 2009-2010 NHIS data matched with the 2010-2011 MEPS files along with additional paradata. Non-response rates are analyzed by standard demographic categories and MEPS response categories.

Key Words: MEPS, NHIS, nonresponse

1. Introduction

Because nonresponse rates were increasing for the MEPS Household component (MEPS-HC), an incentive experiment was conducted with the MEPS panel that began in 2008 (Panel 13) for the entire 5 rounds of MEPS. Incentives of \$30, \$50, and \$70 were tested. Because of favorable results, OMB approved an increase in the incentive from \$30 that MEPS had been using to \$50 starting with the Panel that began in 2011 (Panel 16).

For this paper we compare the MEPS-HC nonresponse rates for the Panel that began in 2011 (Panel 16) that had the increased incentives, with the panel that began in 2010 before the incentive increase. We analyze characteristics of people who could be affected by the increased incentive using NHIS variables. Since NHIS is a sampling frame for MEPS, there is a wealth of information about the sampling frame for both MEPS responders and nonresponders. Our research questions are: did nonresponse rates change from 2010 to 2011 overall and for subgroups; and did the significance of predictive variables for nonresponse vs. a reference group change from 2010 to 2011.

2. Background

The Medical Expenditure Panel Survey (MEPS) is a nationally representative longitudinal survey that collects detailed information on health care utilization and expenditures, health insurance, and health status, as well as on a wide variety of social, demographic, and economic characteristics for the U.S. civilian noninstitutionalized population. MEPS's main sponsor is the Agency for Healthcare Research and Quality. The MEPS has three

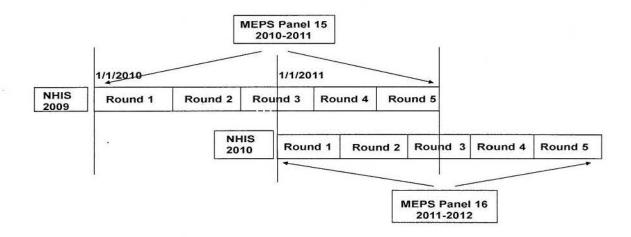
¹ The views in this paper are those of the authors and no official endorsement by the Department of Health and Human Services or the Agency for Healthcare Research and Quality (AHRQ) is intended or should be inferred.

components: the Household component, the Medical Provider component, and the Insurance component. The MEPS-HC collects medical expenditure data at both the person and family levels. The MEPS-HC uses an overlapping panel design in which, for each panel, data are collected covering a two year period by a series of five in-person interviews over the course of two-and-a-half years. Full year public use files are based on Rounds 1-3 data for that year's panel and rounds 3-5 of the previous year's panel.

The MEPS-HC panels of households (Panels 15 and 16, respectively) are subsamples of responding households from the prior year of another large ongoing U.S. health survey, the National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The NHIS sampling frame provides a nationally representative sample of the U.S. civilian non-institutional population and reflects an over sampling of Hispanics, Blacks and Asians. In addition in 2010 and 2011 MEPS oversampled Asians, Hispanics, and Blacks.

Figure 1demonstrates MEPS overlapping panel design for MEPS Panels 15 that began in 2010 that had an incentive of \$30 and Panel 16 that began in 2011 and with an incentive of \$50. The 2010 MEPS uses the 2009 NHIS as its sampling frame. The 2009 NHIS is fielded in 2009 and the 2010 MEPS is fielded in 2010, 2011 and part of 2012. The 5 interviews in MEPS cover the 2-year reference period of 2010 and 2011. The overlapping panel design demonstrates that the second year of the 2010 MEPS panel and the first year of the 2011 MEPS panel both cover the 2011 period.

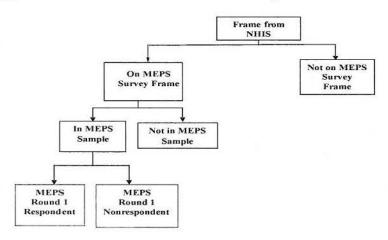
Figure 1: MEPS-HC Overlapping Panel Design with NHIS



3. Methods/Data

The MEPS sampling frame is based on NHIS responding households (Figure 2). The MEPS sample is selected from its sampling frame, and the remaining households in the MEPS sampling frame are not in the MEPS sample. Once MEPS has its sampled households, then there are the MEPS Round 1 responding households versus the MEPS Round 1 nonresponding households. For our study we focused on Round 1 nonresponse and measured Round 1 nonresponse at a person level to be sampled persons who were in round 1 of the current year and did not have a positive weight on the MEPS-HC point-in-time file (PIT) for that year. The MEPS-HC PIT for a given year includes data for Round 1 of that year's panel and data for Round 3 of the previous year's panel. We are using round 1 persons with a positive weight.

Figure 2: MEPS Survey Frame from NHIS



The sizes for the 2010 MEPS-HC cases included in our study are as follows. Of the 88,446 persons on the 2009 NHIS file, 32,171 persons were on the MEPS sampling frame (leaving 56,275 not on the MEPS sampling frame). Of the 32,171 persons on the MEPS sampling frame, 22,906 persons were sampled (and 9,265 were not sampled). Of the 22,906 persons that were sampled, we didn't have complete linkage information for 30 which were excluded. An additional 4,774 persons in Primary Sampling Units (PSUs) that were involved in another MEPS study were also excluded. This resulted in 18,102 sampled persons for the 2010 MEPS-HC used in our analyses, of which 13,201 persons were responding persons, and 4,901 persons were nonreponding.

The sizes for the 2011 MEPS-HC cases included in our study are as follows. Of the 89,976 persons on the 2010 NHIS file, 31,282 persons were on the MEPS sampling frame (leaving 58,694 not on the MEPS sampling frame). A sample of 26,640 persons were selected from the 31,282 persons on the MEPS sampling frame (4,642 were not sampled). Of the 26,640 persons that were sampled, we didn't have complete linkage information for 23 which were excluded. An addition 5,708 persons in PSUs that were involved in another study were excluded. This resulted in 20,909 persons for our study of which 16,380 were responding persons and 4,529 were nonresponding persons.

For our study the 2009 and 2010 NHIS public use files were merged with the MEPS 2010 and 2011 point in time (PIT) files and with the NHIS and MEPS paradata files. The MEPS PIT file has round 1 data for the current panel and round 3 data of the previous panel. We are using the round 1 data of the current panel in both the 2010 and 2011 PIT files. We defined Round 1 nonresponse at a person level to be sampled cases who were not MEPS PIT Round 1 responders. Our estimates are weighted estimates where we start with the NHIS weights and multiply them by the inverse of the MEPS probability of sample selection.

We have three separate analyses in our study. First we calculated and compared person-level nonresponse rates by standard demographic and socioeconomic groups using 2010 and 2011 MEPS linked to the NHIS data.

Next we modelled (again using 2010 and 2011 MEPS linked to the NHIS data) nonresponse as the outcome variable based on a logistic regression model for binary outcomes. The model says that P, the probability of response is related to the covariates X by a logistic regression equation:

$$\log \left[\frac{P}{1-P} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

where P = prob(Y=1|X) = E(Y|X). The regression coefficient estimates ("betas") model effect of covariates on log-odds that Y=1 (Nonrespone).

For the second analyses we modeled nonresponse for each of demographic/socioeconomic variables separately (single predictors of nonresponse); models were run for 2010 and 2011. For our third analyses we modeled nonresponse with all of the socioeconomic variables in the model together (multiple predictors of nonresponse); models were run for 2010 and 2011. Our nonresponse analyses is at the person level. This is a little different from usual nonresponse models that are at the household responding unit level. The demographic, socioeconomic status and paradata variables used are shown in Figure 3.

Demographic	SES	Paradata
Age	Poverty Rate	NHIS completeness status
Sex	Health insurance coverage	
Region	Family health care spending	
Race/Ethnicity		
# kids in RU		
# Adults in RU		
Born in US		
U.S. Citizenship		
Health Status		

Important demographic NHIS variables we used include Age, Sex, Region, Race/Ethnicity (Hispanic, non-Hispanic Black, non-Hispanic White, Other), Number of children (0 versus at least 1 child), Number of adults (1 adult versus 2 or more adults), whether born in the U.S., whether a U.S. citizen, and reported health status where the responses were combined into the two categories excellent/very good/good and fair/poor health. In addition we included poverty rate using 4 categories of percents of the federal poverty level ([0,200), [200,300), [300,400), and 400+); health insurance coverage using 4 categories (Any private, Public only, Other, and Uninsured); and family health care spending using 4 categories (zero, (0,\$2000), [\$2000, \$3000), \$3000+). The NHIS paradata variable of whether or not the NHIS completed interview was partially complete versus fully complete was also included in our model.

4. Results

<u>First analyses:</u> From 2010 to 2011 the MEPS overall nonresponse rate had a relative percent decrease of 17.4 percent from 29.4 percent for 2010 to 24.3 percent for 2011. As shown in Table 1, nonresponse decreased across all subcategories of our variables with only the following exceptions: ages 65-84; ages 85 and over; the Midwest; non-Hispanic other; non-citizens, those in the [200,300) poverty category; and those whose health care spending was greater than zero to \$1,999, or went from \$2,000 to \$2,999.

<u>Second and third analyses</u>: Our second analyses involved modeling nonresponse for each of the variables separately (single predictors of nonresponse) for both 2010 and 2011. Our third analyses modeled nonresponse with the variables combined into the same model (multiple predictors of nonresponse).

Figures 4 and 5 include summaries of the single predictor of nonresponse and the multiple predictor of nonresponse models. One research question was to determine the variable categories whose significance of the odds ratios (ORs) (versus the reference group) changed from 2010 to 2011. These tables only provide those variables in which the significance of the ORs versus the reference group changed from 2010 (before the increased incentive) to 2011 (the first year of the increased incentive) for one of its categories. The ORs comparing a category with the reference group and 95 percent confidence intervals for the ORs are shown. ORs significantly different than 1:00 in a given year are indicated with an asterisk; highlighting indicates the categories in which the ORs' significance level changed from 2011 to 2012.

Variable 2010 2011 V		Variable	2010	2011	
Age	3,5,1		Poverty Rate		
0-17	0.82*(0.74,0.91)	0.77*(0.68),0.86	[0,200)	0.59*(0.50,0.71)	0.54*(0.46,0.62)
18-24	1.05 (0.89,1.23)	1.09 (0.96,1.24)	[200,300)	0.60*(0.48,0.75)	0.66*(0.55,0.80)
25-64	1.00	1.00	[300,400)	0.84 (0.68, 1.02)	0.73*(0.59,0.92)
65-84	0.93 (0.79,1.09)	1.03 (0.90,1.18)	400+	1.00	1.00
85+	1.29 (0.94,1.78)	1.45*(1.13,1.86)			
			Health Ins Cov Any Private	1.00	1.00
U.S. Citizenship Yes	1.00	1.00	Public Only	0.68*(0.57,0.81)	0.55*(0.47,0.65)
No	0.80*(0.66,0.97)	0.84 (0.70,1.01)	Other	0.89 (0.69,1.14)	0.69*(0.53,0.90)
			Not Covered	0.81*(0.70,0.94)	0.68*(0.57,0.80)
Health Status E, VG, G	1.00	1.00			
Fair/Poor	0.84 (70,1.00)	0.79*(0.68,0,93)			

As shown in Figure 4, the age, U.S. Citizenship, Health Status, Poverty Rate, and Health Insurance Coverage had categories whose significance with respect to the reference group for the single predictors of nonresponse models changed from 2010 to 2011. In 2010, persons ages 85 and over were not significantly different than those ages 25-64 (the reference group) in being a nonresponder. In 2011, however, those ages 85 and over were more likely to be a nonresponder than those ages 25-64.

In 2010 Noncitizens were less likely than Citizens (the reference group) to be nonresponders, but in 2011 they are not significantly different than Citizens in being nonresponders.

In 2010 those reported to be in fair or poor (F/P) health were not significantly different than those reported to be in Excellent/Very Good/Good (E/VG/G) health. This changed in 2011 where those in F/P health were less likely to be a nonresponder than those in E/VG/G health.

In 2010, those in the category 300 to less than 400 percent of the federal poverty level were not significantly different than those in the highest level (the reference group) in being a nonresponder, but in 2011 they are less likely than those in the highest level to be a nonresponder.

Persons with Other insurance changed from being not significantly different than those with Any Private (the reference group) in 2010 to in 2011 being less likely to be a nonresponder than those with Any Private.

Figure 5 below provides the summary of results for the multiple predictors model. As shown in Figure 5, Age, Region, Health Insurance Coverage, and Health Care Spending had categories whose significance with respect to the reference group changed from 2010 to 2011 after adjusting for the other variables in the model.

Variable	2010	2011	Variable	2010	2011
Age 0-17	1.02 (0.93,1.12)	1.04 (0.93,1.15)	Health Ins Cov Any Private	1.00	1.00
18-24	1.17 (1.00,1.36)	1.22*(1.06,1.41)	Public Only	0.89 (0.72,1.10)	0.67*(0.55,0.81)
25-64	1.00	1.00	Other	1.21 (0.90,1.63)	0.86 (0.63,1.17)
65-84	0.93 (0.78,1.12)	1.05 (0.89,1.25)	Not Covered	1.05 (0.89,1.23)	0.83 (0.68,1.02)
85+	1.23 (0.85,1.78)	1.56*(1.18,2.08)			

Region			Health Care Spending		
Northeast	1.08 (0.81,1.45)	1.11 (0.83,1.47)	Zero	1.19 (0.93,1.52)	1.04 (0.83,1.30)
Midwest	0.86 (0.65,1.14)	1.08 (0.81,1.44)	(0,\$2000)	1.03 (0.77,1.39)	1.23 (0.92,1.66)
South	1.19 (0.96,1.48)	1.35*(1.06,1.70)	[\$2000,\$3000)	1.00	1.00
West	1.00	1.00	\$3000+	2.04*(1.36,3.06)	1.27 (0.84, 1.91)

After adjusting for the other variables in the model, the odds ratio for persons ages 18-24 and ages 85 and over versus those ages 25-64 (the reference group) went from being not significant in 2010 to having odds ratio in 2011 significantly greater than one. That is, in 2011, persons ages 18-24 and ages 85 and over were more likely to be nonresponders than those ages 25-64. None of the other regions were significantly different than the West (the reference group) in being a nonresponder in 2010. However, the South, which was not significantly different than the West in being a nonresponder in 2010, was more likely to be a nonresponder than the West in 2011.

In 2010, none of the other health insurance categories were significantly different than those with Any private insurance coverage (the reference group). But in 2011, those with Public only were less likely than those with Any private to be a nonresponder. In 2010, those with family health care spending of \$3,000 or more were more likely than those with health care spending from \$2,000 to less than \$3,000 (the reference group) to be a nonresponder. But with the increased incentive in 2011, persons with family spending of \$3,000 or more were not significantly different in being a non-responder than those with health care spending from \$2,000 to less than \$3,000.

5. Summary

Our first analyses determined whether nonresponse rates decreased between 2010 (the year before the incentive increase) and 2011 (the year in which incentives increased.) With the increased incentives in 2011, nonresponse rates had a relative percent decrease of 17.4 percent and decreased for almost all of the subcategories of Age, Sex, Region, Race/Ethnicity, # Kids in the responding unit, # of adults in the responding unit, whether born in the United States, U.S. Citizenship status, Health Status, Poverty Rate, Health insurance coverage, Family Health Spending, and whether or not the NHIS completed interview was a partially complete versus fully complete (there were only 8 exceptions). The nonresponse rates did not decrease significantly from 2010 to 2011 for those ages 65-84; those ages 85 and over; those living in the Midwest, those in the non-Hispanic other category, those who were not citizens of the United States, those whose family income was 200 percent to less than 300 percent of the poverty rate, those whose yearly health care spending was greater than zero and less than \$2000, and those whose health care spending was \$2000 or more but less than \$3000.

Our second and third analyses involved modeling nonresponse for each of variables separately (single predictors of nonresponse) for both 2010 and 2011 and then modeling nonresponse with the variables combined into the same model controlling for the other variables (multiple predictors of nonresponse) for both 2010 and 2011.

Figure 6: Summary of Single Prec 2011	dictors of Nonresponse	Models 2010 to
Categories	2010	2011
25-64 (ref grp) vs. 85+	n.s.	lower
US citizens (ref grp) vs. non-citizens	higher	n.s.
Fair/poor health vs. E/VG/G (ref grp)	n.s.	lower
[300,400) vs. 400+ (ref grp)	n.s.	lower
Other HI vs. Any private (ref grp)	n.s.	lower

When the variables were modelled separately we found 5 categories whose significance with respect to the reference group changed from 2010 to 2011 (Figure 6). Those ages 25-64 versus those ages 85 and over were not significantly different in being nonresponders in 2010, but in 2011 they were less likely to be nonresponders. United States citizens were more likely than non-citizens to be nonresponders in 2010 but were not significantly different than non-citizens in 2011. Those with fair or poor health were not significantly different in being a nonresponder in 2010 than those in excellent/very good/ good health status but were less likely to be nonresponsers in 2011. Those with federal poverty level of 300 percent to less than 400 percent of the federal poverty level were not significantly different in being a nonresponder in 2010 than those in the highest group of 400 percent or more of the federal poverty level; in 2011 those in the [300,400) federal poverty level group were less likely than those in the highest group of 400 percent or more of the federal poverty level. Those with Other health insurance versus Any private insurance were not significantly different in being nonresponders in 2010, but in 2011 they were less likely to be nonresponders.

Figure 7: Summary of Multiple P to 2011	redictors of Nonrespon	ise Models 2010
Categories	2010	2011
25-64 (ref grp) vs. 18-24	n.s.	Lower
25-64 (ref grp) vs. 85+	n.s.	Lower
West (ref grp) vs. South	n.s.	Lower
Public only vs. Any private (ref grp)	n.s.	Lower
\$3000+ vs. [\$2000,\$3000)	Higher	n.s.

After controlling for the other variables the significance of the predictive variable categories changed between 2010 and 2011 for five groups from our multiple predictor model. Persons ages 25-64 were not significantly different than those ages 18-24 years and those ages 85 and over in being a nonresponder in 2010, but in 2011 there were less likely than those ages 85 and over in being a nonresponder. Persons in the West were not significantly different from persons in the South in being a nonresponder in 2010, but in 2011 they were less likely to be a nonresponder than those in the South. Persons with Public only health insurance were not significantly different than those with Any private health insurance in being a nonresponder in 2010 but in 2011 they were less likely to be a nonresponder than those with Any private insurance. People whose health care spending was \$3000 or more were more likely to be a nonresponder than those with health care spending of \$2000 to less than \$3000 in 2010, but in 2011 there was no significant difference in their likelihood of nonresponse.

In conclusion we observed relative percent decreases in nonresponse rates overall and for almost all of the subgroups. We also observed changes in significance of predictive variables for nonresponse versus a reference group from 2010 to 2011. These are preliminary analyses. Additional analyses exploring the effects of the increased incentive on the different dimensions of nonresponse can be further investigated. These dimensions could include levels of cooperation, refusal rates, level of effort as well as possible changes in the quality of the collected data. Although the incentive increase between 2010 and 2011 more than likely explains most of these changes in nonresponse rates, there may be other documented and undocumented changes in field procedures or in the survey climate between 2010 and 2011 that may explain some of these changes. For example in 2011 in another effort to increase the response rates, the round 1 field period was lengthened and this could partly explain some of the observed changes.

Table 1: Estimated weighted nonresponse rates for the Medical Expenditure Panel Survey by demographic characteristics, U.S. 2010-2011^{1,2}

			20	10		2011				100
	w									20
			Non-res	ponders			Non-responders			z-
var_item	resplvl	colper_	secol_2	LB95	UB95	colper_2	secol_2	LB95	UB95	
Age	TOTAL	29.4	0.86	27.8	31.1	24.3	0.72	22.9	25.7	
	0-17 years	26.3	1.32	23.8	28.9	20.3	1.10	18.2	22.5	
	18-24 years	31.4	1.80	27.9	34.9	26.7	1.43	23.9	29.5	
	25-64 years	30.4	0.92	28.5	32.2	25.0	0.79	23.4	26.5	
	65-84 years	28.9	1.57	25.8	31.9	25.6	1.27	23.1	28.1	1100
	85 years and older	36.1	3.68	28.8	43.3	32.6	2.74	27.2	37.9	67883
ex	male	30.4	0.96	28.5	32.3	25.0	0.83	23.4	26.7	
	female	28.5	0.87	26.8	30.2	23.6	0.70	22.3	25.0	
tegion	Northeast	32.9	2.37	28.3	37.6	24.6	1.89	20.9	28.3	
	Midwest	26.0	1.92	22.2	29.8	23.2	1.60	20.1	26.4	1.42
	South	30.9	1.28	28.4	33.4	26.0	1.02	24.0		
	West	28.1		25.0		22.2	1.64	18.9		
tace/Ethnicity	1: Hispan	24.3		21.3		18.0	0.91	16.2		
	2: NH whi	32.1		29.9		26.8	0.91	25.0		
	3: NH Bla	20.3		17.5	the second law to sell and	16.5	1.24	14.1	And the second second	
	4: NH Oth	27.8	2.38	23.2	32.5	26.2	2.39	21.5	30.8	
lumber of children	None	32.3		30.4		27.7	0.92	25.9	29.5	
	One or more	27.0	1.14	24.7		21.4	0.99	19.5	23.4	
lumber of adults	One	27.0		24.4	29.6	22.7	1.07	20.6	24.8	
	Two or more	30.0		28.1		24.7	0.84	23.1		
I.S. Born	Yes	29.6	0.90	27.8	31.3	24.3	0.77	22.8		
	No	27.8		25.2		23.4	1.13	21.2		21
itizenship	Yes	29.6		27.9	The state of the s	24.3	0.74	22.9	25.8	
	No	25.1	COLUMN CARROL	21.4		21.3	1.53	18.3	24.2	2536
overty level	GE 0 and LT 200	24.5	the second second second	22.0		19.2	0.88	17.4		
	GE 200 and LT 300	24.7	SERVICE VALUE OF	21.2		22.6	1.54	19.5	25.6	13073
	GE 300 and LT 400	31.4		27.4		24.5	1.85	20.9		-
	GE 400	35.4	1.34	32.8	38.0	30.7	1.24	28.2	33.1	29
lealth care		la l								
pending/year	Zero	29.5		27.5		23.4	0.80	21.9	25.0	bertien
	Gt 0 to less than \$2,000	27.4		22.8		27.4	2.33	22.8	GERMAN STATES OF THE STATES OF	
	GE \$2,000 to less than \$3,000	26.4		22.2	THE STOP STREET, STOP ST.	24.1	1.93	20.3	Calculation and C	ROBER .
	GE \$3,000	47.8		40.0		36.4	4.09	28.4		
lealth status	1: Exc, VG, G	29.7		28.0		24.7	0.76	23.2		N.
7 25	2: Fair, Poor	26.3	1.65	23.0	29.5	20.6	1.21	18.3	23.0	
ully vs Partially ompleted NHIS										
nterview	Fully completed NHIS int.	26.2	0.89	24.5	28.0	20.3	0.70	18.9	21.7	
	Partially completed NHIS int.	40.9	1.69	37.6	44.2	33.4	1.56	30.3	36.4	
lealth insurance	and desired property and the control of the control	59,5446								
overage	Any private	31.4	0.96	29.5	33.3	27.5	0.92	25.7	29.3	
ST 1770	Public only	23.7	1.63	20.5		17.3	1.07	15.2	19.4	
	Other	28.9		23.9		20.7	2.07	16.7		
	Uninsured	27.1	1.49	24.19		20.4	1.31	17.86	23.00	

¹Nonresponse is round 1 nonresponse at the person level and is defined as MEPS sampled persons who do not have a positive weight on the MEPS -HC point-in-time (PIT) file for that year. A30

²For the 2010 file, 30 persons were excluded because of matching issues and 4,774 persons were excluded because they were in Primary Sampling Units that were involved in another MEPS study resulting in 18,102 persons on this file. For the 2011 file, 23 persons were excluded because of matching issues and 5,708 persons were excluded because they were in Primary Sampling Units that were involved in another MEPS study resulting in 20,909 persons on this file.