

## Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

---

Jeanie Lo

---

Date

**Medical Costs in Chronic Fatigue Syndrome (CFS):  
A Sample of Registry-Based Cases**

By

Jeanie Lo

Master of Public Health

Epidemiology

---

Patrick Sullivan, PhD, DVM

Faculty Thesis Advisor

---

Date

---

Jin-Mann Lin, PhD

Thesis Field Advisor

---

Date

**Medical Costs in Chronic Fatigue Syndrome (CFS):  
A Sample of Registry-Based Cases**

By

Jeanie Lo

B.S., B.A.  
University of California, San Diego  
2006

Thesis Committee Chair: Patrick Sullivan, PhD, DVM

An abstract of  
A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of

Master of Public Health  
in Epidemiology  
2012

## **Abstract**

### **Medical Costs in Chronic Fatigue Syndrome (CFS): A Sample of Registry-Based Cases**

By Jeanie Lo

Chronic fatigue syndrome (CFS) is a complex and debilitating chronic illness that affects at least 4 million people in the United States. Evaluating and assessing the economic impact of CFS will prove crucial in shaping future policies for clinical research, health institutions, and education in order to reduce the burden of disease. The objective of this study analysis is to estimate the medical costs among CFS patients identified from a registry pilot study in primary and tertiary healthcare settings as well as CFS patients identified from a self-referred support group between September 2008 and March 2010 in Georgia, USA.

Participants completed a clinical evaluation to confirm CFS diagnosis and to identify other illnesses. Socio-demographic information including economic data and healthcare utilization was also collected. We estimated the effect of CFS on direct medical costs that included inpatient hospitalizations, provider encounter visits, over-the-counter medications, and other health care costs by stratifying on primary factor-referral status ( $n = 35$ ) from Bibb County, Georgia and Macon City, Georgia. Linear regression models using Ordinary Least Squares were employed to adjust medical costs and earnings for confounders (age, sex, race, marital status, education, working status, healthcare coverage, and unmet need).

Provider-referred CFS patients had mean annual direct medical costs of \$2,462 after adjusting for potential confounders. After adjustment, sex was found to be statistically significant in the other health costs category ( $p=0.04$ ). Additionally, unmet need was found to be statistically significant in both the total annual healthcare expenditures and provider encounter visits categories ( $p=0.02$ ,  $p=0.04$ , respectively). Similarly, educational status was found to be statistically significant in both the total annual healthcare expenditures and provider encounter visits categories ( $p=0.03$ ,  $p=0.03$ , respectively).

These study results demonstrate that CFS may lead to considerable increases in medical costs. There is no known cure for CFS; therefore, treatment and management is long-term and the associated costs may be incurred over decades or even a lifetime. Lastly, this study may offer unique insight via the perspective of a clinical registry sample population.

**Medical Costs in Chronic Fatigue Syndrome (CFS):  
A Sample of Registry-Based Cases**

By

Jeanie Lo

B.S., B.A.  
University of California, San Diego  
2006

Thesis Committee Chair: Patrick Sullivan, PhD, DVM

A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of

Master of Public Health  
in Epidemiology  
2012

## **Acknowledgements**

I would like to gratefully acknowledge and thank my committee members: Jin-Mann Lin, PhD and Patrick Sullivan, PhD, DVM. I acquired much knowledge and skills during the development and completion of this project. This thesis would not have been possible without their expertise, support, insight, and encouragement, and for this, I will be forever grateful.

I would also like to thank Hao Tian, PhD for his guidance, support, and patience as well.

Sincere thanks and gratitude also go out to the Centers for Disease Control and Prevention Chronic Fatigue Syndrome program and study participants. Furthermore, I gratefully acknowledge Dr. Elizabeth Unger, Dr. William Reeves, Dr. Elizabeth Maloney, and Dr. Dana Brimmer. Lastly, much thanks to Abt Associates, Inc. particularly Dr. Scott Royal and Ms. Rebecca Devlin.

## Table of Contents

|                          |    |
|--------------------------|----|
| BACKGROUND .....         | 1  |
| METHODS .....            | 5  |
| RESULTS .....            | 14 |
| DISCUSSION .....         | 18 |
| CONCLUSION .....         | 24 |
| REFERENCES .....         | 25 |
| FIGURES .....            | 29 |
| TABLES .....             | 30 |
| ANNOTATED SAS CODE ..... | 37 |

## BACKGROUND

Chronic fatigue syndrome (CFS) is a complex and debilitating illness characterized by unexplained severe fatigue of at least 6 months duration that interferes with daily activities and the presence of additional symptoms such as post-exertional malaise, muscle aches and pain, and concentration problems (1-3). A population-based study in Georgia has shown that at least 4 million people in the United States suffer from CFS which poses a significant public health problem (4). Average duration of the illness is 7 years, although some individuals may suffer from it throughout their lifetime (5). Currently, the cause of CFS is still a mystery to many scientists and researchers. A single cause for CFS has still not been identified which suggests that CFS may be a multi-factorial disorder (6).

Furthermore, CFS patients, their families, the health care system, and society as a whole incur significant costs associated with the illness. Many of the prominent symptoms of CFS are also common to other illnesses as well. Therefore, diagnosis is complex and requires exclusion of medical and psychiatric conditions that may present correspondingly in other illnesses (1-4). Additionally, the social burden of CFS is poorly recognized although the illness is considered to be one of the most common chronic illnesses among women in the United States (2, 3).

Moreover, there is currently no known cure for CFS nor are there any diagnostic laboratory abnormalities or clinical tests (4). Furthermore, no prevention strategy exists for CFS. Despite the burden CFS imposes on affected individuals and society, less than 20% of those with the illness have been diagnosed and received treatment (3, 7). Hence, treatment and management for CFS is long-term and the associated costs may be incurred over decades or even a lifetime.



Besides the burden and reductions in quality of life imposed on individuals with CFS, the illness greatly limits patients the ability to maintain employment and a steady income (7-9). A previous study had estimated that the average family affected by CFS forgoes approximately \$20,000 in annual earnings and wages (5). Furthermore, a quarter of these individuals with the illness are unemployed or receive disability payments (5). Subsequently, this contributes to an economic burden imposed upon individuals with CFS, their caretakers, the health care system, and society.

Few studies in the United States have assessed the economic impact of CFS. The first study was based on a 4-year longitudinal surveillance study regarding CFS between 1997 to 2000 in Sedgwick County, Kansas that also included Wichita and other unincorporated areas (3). Baseline data were collected during the first year, with follow-up every 12 months until 2000. In 2003, individuals from the 4-year Wichita surveillance study were invited to participate in a clinical evaluation (10). Based on the limited economic cost data collected in the Wichita study, Reynolds et al. estimated the economic impact of CFS in terms of productive loss by employing microsimulation models on the 1997 Wichita baseline data (5). The authors found that the economic burden of CFS was an estimated \$20,000 per patient that amounted to an annual productivity loss of \$9.1 billion in the United States (5).

The second study estimated the economic burden based on two Chicago samples. The first sample was from the community while the second sample was from tertiary healthcare settings. In an epidemiological study of Myalgic Encephalomyelitis (ME/CFS) between 1995 and 1998, Jason et al. estimated the economic impact of ME/CFS (11). However, this study only contained data on direct costs. In order to estimate the total economic costs to society in the United States, this study summed the indirect cost estimates from Reynolds et al.'s study to obtain the combined economic cost estimate of direct and

indirect costs (5, 11). Thus, the authors demonstrated that the mean, direct costs attributed to ME/CFS in Chicago were \$2,342 per patient (11). From the summation described above, the authors estimated that the annual direct cost of ME/CFS to society in the United States was approximately \$2 billion (11).

Lastly, in a case-control study that was conducted between 2004 and 2005, Lin et al. demonstrated the mean, annual direct costs associated with CFS in Georgia were \$3,286 after adjusting for confounding factors (12). Participants in this study underwent a clinical evaluation to confirm CFS diagnosis, and socioeconomic data were also collected for these individuals. After adjusting for confounding factors, Lin et al. also showed that CFS accounted for \$8,554 in lost earnings annually (12).

Furthermore, there have been several studies that have evaluated the economic burden of CFS in a specialist-based study setting. One recent, cross-sectional study from the United Kingdom by Collin et al. demonstrated the mean productivity loss per person attributed to ME/CFS was £22,684 (13). Patient-level data from specialist CFS/ME services in Bristol (Frenchay), Wells (Somerset), Leeds (Leeds & West Yorkshire), Barts and The London, Epsom and St. Helier (South West London and Surrey) in the United Kingdom were analyzed between 2006 and 2010 to estimate the cumulative loss of earnings due to discontinuation of employment (13). Furthermore, the estimated annual productivity cost to the United Kingdom economy was approximately £100 million (13). In 2009, another British study evaluated the financial and psychological impacts on mothers of children with CFS/ME (14). This study used inventories and questionnaires completed from the mothers whose children attended regional special CFS/ME services in Southwest England. Ultimately, this study estimated that the mean income loss attributed to CFS was £247 per

month while the mean, monthly additional expenditures in families with children with CFS/ME was £206 per month (14).

In addition, a study from the Netherlands evaluated the cost of illness and well-being in patients with fibromyalgia (FM) using cost-utility data from FM patients in a randomized, clinical trial (15). This study found that the estimated average, annual total direct costs per FM patient was €5,241 (15). Thus, CFS and other fatiguing illnesses are proving to be a global, public health issue.

Most of the aforementioned CFS studies have consisted mainly of analyzing CFS cases in population-based, specialist services-based, or randomized clinical trial settings. However, no previous study has combined CFS cases identified from a registry-based study into primary and tertiary healthcare settings along with a self-referred support group to assess the economic impact of CFS. To the best of our knowledge, this is the first study in the United States that employed a sample from CFS cases identified from a registry-based study via the combined method of provider- and self-referral from a self-support group. Therefore, the objective of this study is to estimate the medical costs among CFS patients identified from a registry pilot study in primary and tertiary healthcare settings as well as CFS patients identified from a self-referred support group.

## METHODS

Data for this analysis were from the Centers for Disease Control & Prevention (CDC) Registry of Unexplained Fatiguing Illnesses and Chronic Fatigue Syndrome: A Pilot Study. The field data collection was managed by Abt Associates, Inc., and the Institutional Review Board (IRB) of CDC and Abt both approved the study. Since this study analysis did not employ any original identifiers nor was there access to a linking file, this did not qualify as human subject research. Therefore, Emory IRB approval was not required. All participants gave informed consent and were assigned a unique identification number (study ID). All data collected on each participant were labeled with the corresponding study ID number. Personal identifiers such as names, social security numbers, phone numbers, and small geographic information were not recorded in the coded data sets. Only coded datasets were delivered to the CDC. No identification information was provided to the CDC. The pilot Registry was conducted in Bibb County, Georgia from September 1, 2008 to March 20, 2010. This included a 6-month provider recruitment phase followed by a full year of provider-based possible-case referrals to our clinic where patients were clinically evaluated. In December 2009, the Registry expanded to allow members of self-referral possible-cases from a local CFS support group in Macon City. All the amendments were approved by CDC's IRB as well as Abt's IRB.

### *Overview of Study Design*

#### Provider recruitment

In this pilot Registry study, providers were defined as both traditional medical providers (e.g. physicians and nurse practitioners) as well as complementary alternative medical providers (e.g. acupuncturists and physical therapists). Furthermore, the Registry

aimed to target providers who practiced in, or within, 30 miles of Bibb County, Georgia. Primary methods of identifying providers included creating a database of health care providers during the pre-implementation phase, obtaining data from insurance web sites such as United Healthcare, creating a list of the Macon CFS Support Group Providers, etc. The goal was to enroll 600 providers starting on September 1, 2008. Additionally, Registry kick-off packets including information from both Abt Associates and CDC, a provider enrollment verification form, provider frequently asked questions, and a CFS September 10, 2008 Continuing Medical Education (CME) event flier were sent to 1,593 providers (including providers with multiple addresses). Ultimately, a total of 827 providers were enrolled in the pilot Registry consisting of 491 physicians and 336 non-physicians.

#### Subject recruitment and Participants

Upon receipt of the permission-to-be-contacted form (provider-referred patients) or the self-referred phone call (support group members), Registry staff contacted prospective Registry subjects for a screening interview. The screening interview was conducted using a Computer-Assisted Telephone Interview (CATI) from September 1, 2008 through November 30, 2009. In December 2009, given the high eligibility rate, we used a Computer-Assisted Personal Interview (CAPI) to shorten the transition process from the screening interview to the one-day clinical evaluation. The CFS-specific screening questions on the CAPI questions were identical to those on the CATI, but the CAPI was interviewer-administered onsite in the Registry clinic on the same day of the subjects' clinical evaluation. Prior to their clinical evaluation, CAPI subjects received a short screening call prior to clinic scheduling to determine fatigue and symptom inclusion and exclusion criteria. Eligibility criteria for patient referral included unexplained severe fatigue lasting for one month or

longer, one of the following for at least one month (unrefreshing sleep, problems with memory or concentration, or unexplained joint or muscle pain), a body mass index (BMI) of less than 40.0 kg/m<sup>2</sup>, and be aged 12 to 59. However, in November 2009, this age range was expanded to 69 years of age. Exclusionary criteria for patient referral included pregnancy within the last 12 months, stroke with no full recovery, Parkinson's Disease, chronic obstructive pulmonary disease (COPD) or congestive heart failure, insulin-dependent diabetes, uncontrolled diabetes type II (HgbA1c < 9%), anemia, uncontrolled hypo- or hyper-thyroidism, uncontrolled hypertension (blood pressure > 140 / 90 mm Hg), sickle cell anemia, cancer within 5 years (except basal skin) or current chemotherapy, untreated depression, substance abuse problems within the past 2 years, anorexia nervosa or bulimia nervosa within the past 5 years, schizophrenia, bipolar I or II disorder, or dementia, and Hepatitis B or C.

Of the 827 providers enrolled in the pilot Registry, 42 providers referred patients to the Registry with a total of 88 patients. Of these, registry staff could only contact 79 persons because 10 people did not return the permission-to-be-contacted form. As to the self-referral from the local support group, 53 referred themselves to the Registry among the 188 support group members. A total of 131 referred subjects were potentially eligible for the Registry.

The pilot Registry study included screening interviews and subsequent one-day clinical evaluation. During the one-day clinical evaluation, subjects underwent comprehensive medical and psychiatric evaluation in addition to a battery of questionnaires. Of the 78 potentially eligible subjects referred by providers, 12 (15%) were found ineligible at the time of the screening interviews. A total of 104 subjects completed clinical evaluations including 5 adolescents.

### Screening interviews

Initially, within 2 days after receiving patient information, an interviewer contacted the subject by telephone to conduct the CATI in order to screen subjects for the clinical evaluation. During the CATI process, telephone interviewers read questions from a computer screen and entered the responses immediately into a computer database. In the event that the interviewer was unable to reach the subject, a message was left on their answering machine. Subsequent call-backs were made until the interviewer was able to complete the CATI process. Beginning August 2009, CDC determined that it was cost-ineffective to continue with the CATI. Thus, the CAPI was implemented. Instead of a telephone interview, the CAPI involved screening patients during the clinical evaluation with questions regarding fatigue and unwellness symptoms. Based on the screening interviews, subjects were classified as followed:

- 1) CFS-like, characterized by severe fatigue lasting six months or longer that was not alleviated by rest, that caused substantial reduction in occupational, educational, social, or personal activities, and that was accompanied by at least four of the CFS case-defining symptoms.
- 2) Chronically unwell, having chronic ( $\geq$  six months) unwellness with or without fatigue, but not meeting the criteria for CFS.

### Clinical Evaluations

After the CATI interviews were completed, clinical evaluation eligibility was determined. Each subject scheduled for clinical evaluation was sent a packet prior to their appointment containing study information, a sample clinic schedule, an informed consent

document, and questionnaires to complete before arriving at the clinic. Clinical assessment consisted of 5 major areas:

1. Detailed medical history (including gynecological history for female subjects), medication use, and a physical examination by a certified, study physician
2. Saliva, urine, and blood collection for screening medical conditions
3. Psychiatric evaluation
4. Questionnaires regarding symptoms and symptom perception, early and adult life experiences, stress, coping, personality traits, economic impact, and health care utilization
5. Laboratory testing

CFS cases were determined by the Standardized algorithm for the 1994 CFS Case Definition (4, 16), fatigue was measured by the Multidimensional Fatigue Inventory (MFI-20) (17); functional impairment measured by the Medical Outcomes Survey Short Form-36 (SF-36) (18); and accompanying symptoms measured by the CDC Symptoms Inventory (19). In addition to assessing occurrence of the eight specified accompanying CFS symptoms (post-exertional malaise, unrefreshing sleep, impaired memory/concentration, new headaches, sore throat, lymphadenopathy, muscle and joint pain), we also determined the severity cutoff for the CFS symptom summary score. Cases were ascertained if subjects met all the criteria for fatigue, functional impairment, and symptoms. Of 55 subjects without rule-out (exclusionary) conditions of CFS, 37 subjects were classified as CFS and 18 as ISF (Insufficient Symptoms or Fatigue). Figure 1 depicts the participant enrollment and retention flow chart.



## ***Measures***

The objective of this thesis is to quantify the direct costs in CFS cases identified from provider-referrals and self-referrals. Measures included in this analysis were 1) direct costs, 2) primary factor-referral status (provider-referral vs. self-referral), and 3) co-factors such as socio-demographics, healthcare access indicators, and unmet healthcare needs. To evaluate direct costs, a key questionnaire given to the subjects was the economic impact questionnaire (EIQ) that collected information regarding patient health insurance, medical costs, earnings, employment status, and other questions. The EIQ was required to be completed at home before the subject arrived for their clinical evaluation. The healthcare utilization (HU) questionnaire was given to subjects and collected information for all clinic participants such as the number of visits to a health care professional per year, types of healthcare professionals seen, visit purpose, and diagnosis received. We used the responses to two questions in the HU questionnaire to derive unmet healthcare needs. Since the EIQ administered to subjects did not contain all necessary pricing and costing data, the 2005 Medical Expenditure Panel Survey (MEPS) was used (20). Specifically, the 2005 Hospital Inpatient Stays File (HC-094D) was merged with the 2005 Full Year Population Characteristics (HC-097) to impute unit costs for inpatient hospitalizations, healthcare provider encounters, over-the-counter (OTC) costs, and other healthcare costs (20).

### *Primary Outcomes: Direct costs*

We measured direct costs in four healthcare expenditure categories: 1) inpatient hospitalization expenditures (defined as all overnight inpatient hospitalizations that occurred in the past year. Data regarding reasons for hospitalization, duration, and any diagnostic tests or surgical procedures performed were collected.); 2) healthcare provider encounters (defined as all encounters with healthcare providers that occurred in the past six months.

Individuals had to report the number of visits to healthcare providers in separate categories, “Medical Doctor or Osteopath”, “Nurse/Paramedical”, “or Psychiatrist/Psychologist/Counselor”. Participants also had to report any encounters with other types of healthcare providers. Encounters that were part of an overnight hospitalization were excluded.); 3) OTC medications (defined as any non-prescription medication expenses incurred in the past four weeks); and 4) other healthcare costs (defined as any other health-related expenses in the past four weeks). If subjects had expenses, each supply or service and the amount spent was reported. Out-of-pocket (OOP) expenses that were incurred during overnight inpatient hospitalizations and encounters with healthcare providers were excluded, as well as any medication costs already reported in previous questions). We further aggregated the healthcare expenditures of the four categories into the total annual healthcare expenditures by summing the costs of the aforementioned three categories. We compared the incremental burden of CFS between provider-referred and self-referred cases.

#### Socio-demographics, healthcare access indicators, and unmet needs

Covariates identified to be possible correlates to the economic burden and healthcare utilization of CFS included socio-demographic characteristics such as sex, race, age, current marital status, education, working status, healthcare coverage, and unmet need (12). Initially, race was stratified into three categories (“White”, “Black”, or “Other”). In later parts of the analysis, race was re-defined as “White” and “Non-white” (where “Non-white” was created by collapsing the “Black” and “Other” races together). This was done to account for the small sample size and to increase study robustness. Similarly, age was initially stratified into 10-year age groups and was later collapsed into “18-39”, “40-59”, and “60-69” categories for

the same reason. Current marital status was defined to be either “Married” or “Not married”. “Not married” encompassed a wide variety of categories such as “not married but living with partner”, “separated”, “divorced”, “widowed”, or “never been married”. Subjects’ educational status was stratified by “college degree or higher” and “less than college degree”. “College degree” was defined as possession of a four-year college degree and included categories such as “graduated 4 year college”, “part graduate/professional school”, and “completed graduate/professional school”. Additionally, “less than college degree” included categories such as “Grade 6 or less”, “Grade 7 to 12 (without graduating high school)”, “graduated high school or high school equivalent”, “part college”, or “graduated 2 year college”. Working status was stratified by “full-time”, “part-time”, and “not currently working” where “full-time” denoted working at least 30 hours per week, and “part-time” was defined as working less than 30 hours per week. The number of healthcare coverage types that a subject possessed was stratified by “two or more”, “one”, or “none”. Lastly, unmet need was a binary variable that posed the question, “*During the past 12 months, have you wanted to or thought that you should consult a healthcare professional, but did not?*”

### ***Statistical Analysis***

Statistical analyses described associations of the direct and total costs variables (dependent variables) and socio-demographic characteristics (independent variables). These analyses only concerned 35 subjects with CFS aged 18 or older. Initially, bivariate analysis was conducted to determine frequencies and crude associations between each of the independent variables and the primary exposure variables. Descriptive statistics were reported as counts and percentages with p-values. Additionally, two-sample t-tests and chi-square tests were conducted in examining the bivariate association of referral groups with

continuous and categorical covariates. Multivariate linear regression analysis with Ordinary Least Square (OLS) estimates was performed to statistically evaluate both unadjusted and adjusted effects of CFS on the three healthcare expenditure categories, in addition to total healthcare expenditures. The primary dependent variable under consideration was factor-referral status (provider-referred or self-referred). Furthermore, the models also contained covariates to adjust for potential confounding; sex, race, age, marital status, education, working status, healthcare coverage, and unmet need. These covariates were chosen as potential socio-demographic confounders due to their association with CFS and healthcare costs. Previous studies have demonstrated that CFS is associated with socio-demographics such as sex or gender (2, 4, 21), as well as race or ethnicity (2, 22). Unadjusted and adjusted estimates associated with direct medical costs and other covariates, along with standard errors (SE) were computed. All data analyses were completed by using SAS 9.2 (SAS Institute, Inc, Cary, NC). The significance for all the tests was set at two-sided significance of 0.05.

## RESULTS

Table 1 summarizes descriptive statistics for the 35 CFS Registry subjects. Overall, the subjects had a mean age of 48 years (standard deviation [SD] 12 years) and were predominantly female (89%), white (89%), and married (74%). In terms of education, 63% of the sample had less than a four-year college degree. Furthermore, 43% of subjects were employed either full-or part-time. Ninety-four percent of the subjects reported having health insurance. Of the insured, 11% had multiple types of healthcare coverage. Additionally, approximately 46% of the study subjects reported having worked at all in the last 4 weeks with average earnings amounting to \$12,821 annually. Lastly, 60% of the sample did not have any unmet need and 71% ( $n = 23$ ) of the subjects was referred by providers.

Moreover, there was no factor statistically different between provider-referred and self-referred CFS subjects except for age (Table 1,  $p = 0.04$ ). All other factors were statistically insignificant. Similar to the overall results, sex and racial proportions in both groups were predominantly female and white (Table 1). In the provider-referred group, most subjects fell into the age 30-39 range (80%) while most subjects in the self-referred group tended to be in the 60-69 age range (75%). The provider-referred group had lower mean age than the self-referred group (45 vs. 54 years,  $p = 0.04$ ). Additionally, 65% of the provider-referred subjects were married compared to 35% of the self-referred subjects ( $p = 1.00$ ). The provider-referred group had a higher level of educational attainment compared to the self-referred group (77% vs. 23% possessing a college degree or higher). Eighty percent of the provider-referred group reported working full-time compared to 20% of the self-referred group working full-time ( $p = 0.60$ ). Furthermore, the percentage of subjects in the provider-referred group who worked at all in the past 4 weeks was slightly over three times

the corresponding percentage in the self-referred group ( $p = 0.29$ ). Consequently, the self-referred group had lower mean annual earnings than the provider-referred group (\$7,545 vs. \$15,573,  $p = 0.25$ ). Both referral types had more people who reported having their healthcare needs met.

Table 2 summarizes the mean annual healthcare expenditures for CFS Registry patients stratified by provider type. We examined healthcare expenditures by four categories: inpatient hospitalization, ambulatory provider visits, OTC medications, and other healthcare costs. Mean costs for inpatient hospitalization were almost four-fold higher for self-referred patients versus provider-referred patients. However, mean costs for ambulatory provider visits were \$11,213 for provider-referred patients compared to \$7,900 for self-referred patients. Furthermore, self-referred patients tended to spend slightly more on OTC medications compared to provider-referred patients. Lastly, self-referred patients spent almost 2.5 times more on other healthcare costs compared to provider-referred patients. All figures in Table 2 are provided in 2005 United States Dollar (USD) amounts.

Table 3 summarizes unadjusted results from Linear OLS for bivariate association with the four healthcare expenditure categories. Compared to self-referred CFS cases, the incremental costs of provider-referred CFS cases for the total expenditures were estimated to be \$2,583. Furthermore, CFS subjects with unmet need spent \$8,186 more in total expenditures than those who did not have any unmet need ( $p = 0.02$ ). In the provider encounters category, there was a statistically significant difference between CFS subjects who had a college degree or higher compared to those who had less than a college degree. On average, CFS subjects with a college degree or higher spent \$8,095 more in provider encounter healthcare than CFS subjects who had less than a college degree. CFS subjects with one or more types of healthcare coverage spent \$9,080 more than CFS subjects who

did not have any healthcare coverage. In the OTC medications category, there was a statistically significant difference between subjects aged 60-69 compared to subjects aged 40-59. Moreover, subjects aged 60-69 spent \$294 more on OTC medications compared to subjects aged 40-59. Additionally, sex was statistically significant associated with other healthcare expenditures; female CFS subjects spent \$321 less compared to male CFS subjects.

Table 4 summarizes the results from the multiple linear regression analysis on the Annual Healthcare Expenditures by provider type, adjusting for sex, race, age, marital status, education, working status, healthcare coverage, and unmet need. When compared to self-referred CFS subjects, the adjusted incremental total expenditures of provider-referred CFS subjects were estimated to be \$2,462. Furthermore, the adjusted incremental provider encounter expenditures for provider-referred CFS subjects were estimated to be \$2,520. Additionally, CFS subjects who had a college degree or higher had \$8,932 more in total healthcare expenditures than those who had less than a college degree. Moreover, CFS subjects who had a college degree or higher spent \$9,572 more in provider encounter healthcare compared to those who had less than a college degree. Subsequently, there were statistically significant differences in both of these cost categories for CFS subjects who had a college degree compared to those who had less than a college degree. Furthermore, CFS subjects with unmet need had \$9,721 more in total expenditures and \$8,455 more in provider encounter expenditures than those who did not have any unmet need. There were statistically significant differences in both total expenditures and provider encounter expenditures for CFS subjects with unmet need compared to those who did not have any unmet need. Lastly, female sex was another statistically significant factor given that female

CFS subjects spent, on average, an estimated \$382 less in other health costs expenditures compared to male CFS subjects.



## DISCUSSION

This study showed that provider-referred subjects with CFS had considerably increased medical costs and greater unmet need compared to self-referred subjects with CFS. Furthermore, this analysis found that the unadjusted, total incremental costs attributed to CFS for provider-referred patients were \$2,583 when compared to self-referred patients. After adjusting for socioeconomic factors and healthcare coverage, this analysis estimated that the incremental total costs attributed to CFS were \$2,462 for provider-referred patients when compared to self-referred patients. The estimates of the direct medical costs found in this analysis are consistent with the findings from previous studies. For example, Lin et al. estimated that direct medical costs were \$3,226 in Georgia while Jason et al. estimated the corresponding costs were \$2,342 in Chicago (11, 12). One potential reason for the slight differences may be due to the fact that Lin et al. included prescription medications in their direct costs analyses whereas this analysis did not include prescription medications (12). Furthermore, this analysis had four distinct categories for estimating direct costs whereas Jason et al.'s study considered current medication, medical tests, and medical office visit prices (11).

Furthermore, these estimates are consistent with findings from other chronic illnesses. To illustrate, Shenolikar et al. estimated that the direct costs of asthma among American working adults was \$3,762 in 2011 (23). Additionally, Lee KW estimated that the direct medical costs for patients with Type 2 diabetes mellitus (without diabetic complications) in Korea were \$1,184,563 won (approximately \$1,184 USD) (24). In a recent study, Wagner E. estimated that the direct medical costs for patients with chronic, non-specific low back pain in Austria were €1443 per year (equivalent to \$1,890 USD) (25). Therefore, not only will evaluating the economic impact of CFS prove crucial to related

fatiguing illnesses, but it can also provide a framework model for estimating the costs of other chronic illnesses as well.

In both the unadjusted and adjusted models, sex in the other health costs category was found to be statistically significant. Potentially, this might be due to the fact that females spend substantially more than males on provider encounter expenditures and inpatient hospitalization expenditures and subsequently would not need to spend as much on other health costs. A study conducted by Bastida et al. evaluated the financial impact of diabetes on adult employment and earnings in a Mexican-American study population (26). Taking sex into consideration, this study found that diabetes led to lower productivity and earnings for women more so compared to men (26). Extrapolating to CFS, an illness that is predominant in women, these findings will prove to be important in the future as researchers strive to better assess the financial burden of CFS and related fatiguing illnesses. Additionally, another explanation for this finding could be due to the fact that other health costs is a very broad category that groups other miscellaneous health care related costs that do not fall into the other three categories.

Another notable finding included the fact that having a college degree or higher was statistically significant in the unadjusted model for provider encounter expenditures, and this statistical significance was also reflected in both the total expenditures and the provider encounter expenditures in the adjusted model. CFS patients in this study with a college degree or higher tend to spend significantly more on annual healthcare expenditures when compared to those who had less than a college degree. Furthermore, provider encounters make up a significant portion of total healthcare expenditures when evaluating educational status alone. An explanation for this finding could be that individuals who are college-educated tend to obtain higher earnings and are more likely to obtain health insurance

through their employers compared to those without a college degree (27, 28). Therefore, college-educated individuals have potentially increased access to care and may afford to visit their providers more often than individuals without a college degree.

A major focus of this study analysis concerned healthcare utilization and unmet need. In the adjusted model, having unmet need was found to be statistically significant in both the total expenditures and provider encounter expenditures categories. Hence, unmet need played an important role in influencing total annual healthcare expenditures. Since our study population is derived from a clinical registry, it is reasonable that our study participants would still have unmet need than when, for example, compared to a population-based study (29). Furthermore, having higher unmet needs could also be indicative of having a more severe level of CFS. Although an individual with CFS may possess health insurance coverage, there could still exist barriers to healthcare utilization such as busy work schedules, inability to travel to a physician's office, lack of knowledge regarding clinic availability, or general CFS-associated impairment (30, 31).

Lastly, being aged 60-69 was statistically significant in the OTC medications category in the unadjusted model, but not in the adjusted model. On average, self-referred CFS subjects were significantly older than provider-referred CFS subjects. A majority of the self-referred CFS subjects were in the 60-69 age grouping. Individuals in this age group are most likely retired and might be living on a lower, fixed income. Therefore, they might be more likely than younger individuals to self-manage their CFS by spending more on OTC medications than on provider encounters. Furthermore, an individual aged 60-69 might suffer from increased infirmities and ailments compared to younger individuals. Hence, it is reasonable that individuals in the older age groups might spend substantially more on annual healthcare expenditures than individuals in the younger age groups. Additionally, since older

individuals comprise a large part of the self-referred subject study population, age can be considered to be a key influence on annual healthcare expenditures (32-34).

In particular, provider encounters appear to be the main force driving total annual healthcare expenditures. On average, provider-referred CFS subjects tend to spend an estimated 1.5 times more on ambulatory provider visits compared to self-referred subjects. However, inpatient hospitalization mean costs were almost four-fold higher for self-referred subjects compared to provider-referred subjects. An explanation for these findings concerns the issue of healthcare utilization and access to care when taking into consideration the socio-demographic backgrounds of provider-referred patients versus self-referred patients. From this sample population, a higher percentage of provider-referred patients had multiple types of healthcare coverage when compared to self-referred patients. Furthermore, inpatient hospitalization costs may involve potentially life-threatening situations and thus, an individual, would feel more obliged to check into a hospital, regardless of cost. However, provider encounters usually encompass less immediate and urgent medical issues; therefore, self-referred patients and/or patients with decreased access to care may not feel as compelled to visit their providers on a regular basis. Hence, these findings demonstrate the influence of healthcare access and healthcare utilization on annual healthcare expenditures.

Moreover, there is an extensive and increasing body of literature that is evaluating the impact of illness on earnings and productivity, in addition to medical costs. Previous studies have demonstrated that illness in general can lead to deteriorating health in the long-term and contribute to lower lifetime earnings (35-38). To reiterate, the direct medical costs associated with CFS in this study analysis appear to be driven by both total annual healthcare expenditures and provider encounter visits. Future studies should consider improvements in CFS diagnoses and primary/tertiary care for more efficient healthcare resource allocation.

### *Strengths and Limitations*

One of the important strengths of this study is using a sample identified from a registry-based, epidemiological study including regular healthcare users to physician clinics and those with additional support from a group for subjects with CFS or FM diagnosis. Furthermore, this study analysis replicates similar methods from Lin et al.'s study to validate and compare results from a registry-based sample versus a population-based sample (12). Using a registry-based sample may be beneficial since it can offer a unique perspective regarding CFS-related characteristics in order to develop potential intervention strategies for CFS. A second strength of this study is that data was collected according to standardized procedures (e.g. structured questionnaires/instruments). Third, an ample amount of economic and healthcare utilization data existed for us to conduct our analyses. Fourth, this study benefited from a thorough study design that conducted CATI and CAPI to efficiently screen individuals for CFS prior to clinical evaluation.

One limitation of this study analysis is that the sample may not capture non-English speaking individuals. For example, the study was initially screened using a random-digit-dial, computer-assisted phone screening in English that thereby excluded non-English speaking individuals. Previous studies have demonstrated that Hispanics and Blacks have similar prevalence of CFS as whites; however, they may not be fully represented in this sample. Second, another limitation is the potential recall bias on the HU and EI questionnaires due to self-reported data. However, a shorter recall period was used for some categories of healthcare expenditures. For example, the respondents were asked for outpatient encounters with healthcare providers in the past 6 months, prescription medications filled in the last 4 weeks, and other medical supplies and services obtained in the past 4 weeks. Furthermore, there was a lack of data on specific dates for healthcare utilization which did not allow this

study analysis to assess the effect of time on healthcare utilization and economic impact. While the sample size in this study analysis ( $n = 35$ ) is relatively small compared to other CFS studies, this study analysis may offer a different perception of the economic impact of CFS with potential implications for future studies. Third, the sampling frame only included the non-institutionalized population aged 18-69. Finally, this study analysis included data collected only on Bibb County and Macon City residents and prevented us from the generalization for the United States. Therefore, the actual individual and societal burden of CFS in the United States could be quite larger than estimated.

## CONCLUSION

In conclusion, individuals who suffer from CFS carry an economic burden in terms of direct medical costs. Furthermore, these individuals may require more healthcare utilization and have more unmet need than other populations. This study demonstrates that provider-referred CFS subjects tend to spend more on total, annual healthcare expenditures when compared to self-referred CFS subjects. Taking our study limitations into consideration, we cannot a-priori generalize these findings beyond the study population. However, these results offer an initial framework for estimating the annual healthcare expenditures for individuals who suffer from CFS. Additionally, these findings offer insight from the perspective of a clinical registry sample population. Therefore, this pilot study can be viewed as a first step towards implementing analyses to assess direct medical costs, healthcare utilization, and unmet need by influencing future policies in clinical health institutions, research, and education.

## REFERENCES

1. Fukuda K, Straus SE, Hickie I, et al. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med* 1994;121(12):953-9.
2. Jason LA, Richman JA, Rademaker AW, et al. A community-based study of chronic fatigue syndrome. *Arch Intern Med* 1999;159(18):2129-37.
3. Reyes M, Nisenbaum R, Hoaglin DC, et al. Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas. *Arch Intern Med* 2003;163(13):1530-6.
4. Reeves WC, Jones JF, Maloney E, et al. Prevalence of chronic fatigue syndrome in metropolitan, urban, and rural Georgia. *Popul Health Metr* 2007;5:5.
5. Reynolds KJ, Vernon SD, Bouchery E, et al. The economic impact of chronic fatigue syndrome. *Cost Eff Resour Alloc* 2004;2(1):4.
6. Prevention CfDC. CDC- General Information - Chronic Fatigue Syndrome (CFS). Atlanta, GA; 2010. (<http://www.cdc.gov/cfs/general/index.html>). (Accessed February 21 2012).
7. Bombardier CH, Buchwald D. Chronic fatigue, chronic fatigue syndrome, and fibromyalgia. Disability and health-care use. *Med Care* 1996;34(9):924-30.
8. Taylor RR, O'Brien J, Kielhofner G, et al. The occupational and quality of life consequences of chronic fatigue syndrome/myalgic encephalomyelitis in young people. *Br J Occup Ther* 2010;73(11):524-30.
9. Lloyd AR, Pender H. The economic impact of chronic fatigue syndrome. *Med J Aust* 1992;157(9):599-601.
10. National Center for Infectious Diseases CfDCP. Clinical Assessment of Subjects with Chronic Fatigue Syndrome and Other Fatiguing Illnesses in Wichita. Atlanta,



- GA; 2011. (<http://www.cdc.gov/cfs/pdf/wichita-data-access/wichita-clinical-study-overview.pdf>). (Accessed February 20 2012).
11. Jason LA, Benton MC, Valentine L, et al. The economic impact of ME/CFS: individual and societal costs. *Dyn Med* 2008;7:6.
  12. Lin JM, Resch SC, Brimmer DJ, et al. The economic impact of chronic fatigue syndrome in Georgia: direct and indirect costs. *Cost Eff Resour Alloc* 2011;9(1):1.
  13. Collin SM, Crawley E, May MT, et al. The impact of CFS/ME on employment and productivity in the UK: a cross-sectional study based on the CFS/ME national outcomes database. *BMC Health Serv Res* 2011;11:217.
  14. Missen A, Hollingworth W, Eaton N, et al. The financial and psychological impacts on mothers of children with chronic fatigue syndrome (CFS/ME). *Child Care Health Dev* 2011.
  15. Boonen A, van den Heuvel R, van Tubergen A, et al. Large differences in cost of illness and wellbeing between patients with fibromyalgia, chronic low back pain, or ankylosing spondylitis. *Ann Rheum Dis* 2005;64(3):396-402.
  16. Reeves WC, Wagner D, Nisenbaum R, et al. Chronic fatigue syndrome--a clinically empirical approach to its definition and study. *BMC Med* 2005;3:19.
  17. Smets EM, Garssen B, Bonke B, et al. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;39(3):315-25.
  18. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-33.

19. Wagner D, Nisenbaum R, Heim C, et al. Psychometric properties of the CDC Symptom Inventory for assessment of chronic fatigue syndrome. *Popul Health Metr* 2005;3:8.
20. Quality AfHRA. Medical Expenditure Panel Survey. 2007. (<http://meps.ahrq.gov/mepsweb/>). (Accessed February 21 2012).
21. Evengard B, Jacks A, Pedersen NL, et al. The epidemiology of chronic fatigue in the Swedish Twin Registry. *Psychol Med* 2005;35(9):1317-26.
22. Dinos S, Khoshaba B, Ashby D, et al. A systematic review of chronic fatigue, its syndromes and ethnicity: prevalence, severity, co-morbidity and coping. *Int J Epidemiol* 2009;38(6):1554-70.
23. Shenolikar R, Song X, Anderson JA, et al. Costs of asthma among US working adults. *Am J Manag Care* 2011;17(6):409-16.
24. Lee KW. Costs of diabetes mellitus in Korea. *Diabetes Metab J* 2011;35(6):567-70.
25. Wagner E. [Costs of non-specific low back pain in Austria.]. *Wien Med Wochenschr* 2012.
26. Bastida E, Pagan JA. The impact of diabetes on adult employment and earnings of Mexican Americans: findings from a community based study. *Health Econ* 2002;11(5):403-13.
27. Blaug M. The correlation between education and earnings: What does it signify? *Higher Education* 1972;1(1):53-76.
28. Hadley J. Sicker and poorer--the consequences of being uninsured: a review of the research on the relationship between health insurance, medical care use, health, work, and income. *Med Care Res Rev* 2003;60(2 Suppl):3S-75S; discussion 6S-112S.

29. Lin JM, Brimmer DJ, Boneva RS, et al. Barriers to healthcare utilization in fatiguing illness: a population-based study in Georgia. *BMC Health Serv Res* 2009;9:13.
30. Solomon L, Reeves WC. Factors influencing the diagnosis of chronic fatigue syndrome. *Arch Intern Med* 2004;164(20):2241-5.
31. Thanawala S, Taylor RR. Service utilization, barriers to service access, and coping in adults with chronic fatigue syndrome. *Journal of Chronic Fatigue Syndrome* 2007;14(1):5-23.
32. Banerjee S. Expenditure patterns of older Americans, 2001-2009. *EBRI Issue Brief* 2012(368):1-25.
33. Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Intern Med* 2002;162(20):2269-76.
34. Yang Z, Norton EC, Stearns SC. Longevity and health care expenditures: the real reasons older people spend more. *J Gerontol B Psychol Sci Soc Sci* 2003;58(1):S2-10.
35. Ettner SL. New evidence on the relationship between income and health. *J Health Econ* 1996;15(1):67-85.
36. Imlach Gunasekara F, Carter KN, Liu I, et al. The relationship between income and health using longitudinal data from New Zealand. *J Epidemiol Community Health* 2011.
37. Fritzell J, Neramo M, Lundberg O. The impact of income: assessing the relationship between income and health in Sweden. *Scand J Public Health* 2004;32(1):6-16.
38. Chirikos TN. The relationship between health and labor market status. *Annu Rev Public Health* 1993;14:293-312.

## FIGURES

**Figure 1. Flow chart of study recruitment and enrollment in the Registry of Unexplained Fatiguing Illnesses and Chronic Fatigue Syndrome: A Pilot Study, United States, 2010.**



## TABLES

Table 1. Socio-demographic Characteristics of Registry CFS Subjects

| Characteristics          | Overall CFS<br>(n = 35) | Provider-referred<br>(n = 23) | Self-referred<br>(n = 12) | p-value <sup>a</sup> |
|--------------------------|-------------------------|-------------------------------|---------------------------|----------------------|
|                          | No. (%)                 | No. (%)                       | No. (%)                   |                      |
| <b>Sex</b>               |                         |                               |                           | 1.00                 |
| Female                   | 31 (89)                 | 20 (65)                       | 11 (36)                   |                      |
| Male                     | 4 (11)                  | 3 (75)                        | 1 (25)                    |                      |
| <b>Race</b>              |                         |                               |                           | 0.11                 |
| White                    | 31 (89)                 | 22 (71)                       | 9 (29)                    |                      |
| Black                    | 3 (9)                   | 1 (33)                        | 2 (67)                    |                      |
| Other                    | 1 (3)                   | 0 (0)                         | 1 (100)                   |                      |
| <b>Age groups</b>        |                         |                               |                           | 0.11                 |
| 18 to 29                 | 2 (6)                   | 2 (100)                       | 0 (0)                     |                      |
| 30 to 39                 | 10 (29)                 | 8 (80)                        | 2 (20)                    |                      |
| 40 to 49                 | 5 (14)                  | 4 (80)                        | 1 (20)                    |                      |
| 50 to 59                 | 10 (29)                 | 7 (70)                        | 3 (33)                    |                      |
| 60 to 69                 | 8 (23)                  | 2 (25)                        | 6 (75)                    |                      |
| <b>Age, years</b>        |                         |                               |                           | 0.04 <sup>b</sup>    |
| Mean                     | 48                      | 45                            | 54                        |                      |
| Standard deviation       | 12                      | 12                            | 12                        |                      |
| <b>Marital status</b>    |                         |                               |                           | 1.00                 |
| Married                  | 26 (74)                 | 17 (65)                       | 9 (35)                    |                      |
| Not married              | 9 (26)                  | 6 (67)                        | 3 (33)                    |                      |
| <b>Education</b>         |                         |                               |                           | 0.46                 |
| College degree or higher | 13 (37)                 | 10 (77)                       | 3 (23)                    |                      |
| Less than college degree | 22 (63)                 | 13 (59)                       | 9 (41)                    |                      |

|                                     |          |          |         |      |
|-------------------------------------|----------|----------|---------|------|
| <b>Working status</b>               |          |          |         | 0.60 |
| Full-time                           | 10 (29)  | 8 (80)   | 2 (20)  |      |
| Part-time                           | 5 (14)   | 3 (60)   | 2 (40)  |      |
| Not currently working               | 20 (57)  | 12 (60)  | 8 (40)  |      |
| <b>Multiple healthcare coverage</b> |          |          |         | 1.00 |
| Two or more                         | 4 (11)   | 3 (75)   | 1 (25)  |      |
| One                                 | 29 (83)  | 19 (66)  | 10 (35) |      |
| None                                | 2 (6)    | 1 (50)   | 1 (50)  |      |
| <b>Employment/Earnings</b>          |          |          |         | 0.25 |
| % who worked at all in last 4 weeks | 46       | 34       | 11      |      |
| Annual earnings                     | \$12,821 | \$15,573 | \$7,545 |      |
| <b>Unmet Need</b>                   |          |          |         | 0.72 |
| Yes                                 | 14 (40)  | 10 (44)  | 4 (33)  |      |
| No                                  | 21 (60)  | 13 (57)  | 8 (67)  |      |

<sup>a</sup>Some numbers may not add up to 100 due to rounding.

<sup>b</sup>Two-sided p-value comparing Provider-referred vs. Self-referred. Bolded values indicate significance. P-values are from chi-square for categorical variables and from t-tests for continuous variables.

**Table 2. Mean Annual Healthcare Expenditures of the Sampled Population (in US Dollars, \$2005)**

*Healthcare Expenditure Category*

| <b>Sample</b>         | <b>Inpatient Hospitalization, mean costs</b> | <b>Ambulatory Provider Visits, mean costs</b> | <b>Over-the-counter Medications; mean costs</b> | <b>Other Healthcare Costs; mean costs</b> |
|-----------------------|--|---|---|---|
| <b>CFS – Registry</b> |  |   |   |   |
| Provider-referred     | 258  | 11,213  | 301   | 49  |
| Self-referred         | 988  | 7,900   | 379   | 119                                       |

**Table 3. Unadjusted Estimated Parameters from the Annual Healthcare Expenditures Model (in US Dollars, \$2005)**

Linear Regression using Ordinary Least Squares (OLS) results

| <i>Annual Healthcare Expenditures</i> |                                      |                                      |                                      |  |                                      |
|---------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--|--------------------------------------|
|                                       | <b>Total</b>                         | <b>Inpatient<br/>Hospital</b>        | <b>Provider<br/>Encounters</b>       | <b>Over-the-<br/>counter<br/>Medications</b> | <b>Other<br/>Health Costs</b>        |
| <b>Independent Variables</b>          | <b>Estimate<br/>(Standard Error)</b> | <b>Estimate<br/>(Standard Error)</b> | <b>Estimate<br/>(Standard Error)</b> | <b>Estimate<br/>(Standard Error)</b>         | <b>Estimate<br/>(Standard Error)</b> |
| <b>CFS Registry</b>                   |                                      |                                      |                                      |  |                                      |
| Provider-referred                     | 2,583.29 (3,780.85)                  | -730.26 (790.14)                     | 3,313.55 (3,756.30)                  | -78.47 (130.96)                              | -69.99 (105.72)                      |
| Self-referred <sup>a</sup>            | -                                    | -                                    | -                                    | -  | -                                    |
| <b>Sex</b>                            |                                      |                                      |                                      |  |                                      |
| Female                                | 319.81(5,680.19)                     | 573.68 (1,189.79)                    | -253.87 (5,669.59)                   | -257.59 (191.26)                             | <b>-321.02* (148.62)</b>             |
| Male <sup>a</sup>                     | -                                    | -                                    | -                                    | -  | -                                    |
| <b>Race</b>                           |                                      |                                      |                                      |  |                                      |
| White                                 | 5,035.77 (5,612.42)                  | 573.68 (1,189.79)                    | 4,462.10 (5,616.30)                  | 204.75 (193.19)                              | 82.61 (158.12)                       |
| Non-white <sup>a b</sup>              | -                                    | -                                    | -                                    | -  | -                                    |
| <b>Age groups</b>                     |                                      |                                      |                                      |  |                                      |
| 18-39                                 | 1,290.20 (4,179.06)                  | -395.20 (854.56)                     | 1,685.40 (4,182.19)                  | -3.96 (131.67)                               | -75.40 (114.44)                      |
| 40-59 <sup>a</sup>                    | -                                    | -                                    | -                                    | -  | -                                    |
| 60-69                                 | 2,978.37 (4,723.96)                  | 1,086.80 (965.99)                    | 1,891.57 (4,727.50)                  | <b>293.64* (139.79)</b>                      | 103.35 (129.36)                      |
| <b>Marital status</b>                 |                                      |                                      |                                      |  |                                      |
| Married                               | 2,545.03 (4,111.31)                  | -202.67 (868.44)                     | 2,747.69 (4,099.51)                  | 73.50 (142.43)                               | 98.50 (114.30)                       |
| Unmarried <sup>a</sup>                | -                                    | -                                    | -                                    | -  | -                                    |
| <b>Education</b>                      |                                      |                                      |                                      |  |                                      |
| College degree or higher              | 7,286.81 (3,518.69)                  | -808.36 (773.48)                     | <b>8,095.17*(3,457.12)</b>           | 112.00 (127.87)                              | 185.86 (99.41)                       |
| Less than college degree <sup>a</sup> | -                                    | -                                    | -                                    | -  | -                                    |
| <b>Working status</b>                 |                                      |                                      |                                      |  |                                      |
| Full or part-time <sup>a</sup>        | -                                    | -                                    | -                                    | -  | -                                    |
| Not working                           | 3,612.33 (3,597.48)                  | 889.20 (751.85)                      | 2,723.13 (3,614.19)                  | 218.40 (120.44)                              | 128.05 (99.62)                       |



|                            |                            |                   |                     |                 |                  |
|----------------------------|----------------------------|-------------------|---------------------|-----------------|------------------|
| <b>Healthcare coverage</b> |                            |                   |                     |                 |                  |
| One or more <sup>c</sup>   | 9,080.12 (7,624.02)        | 538.91 (1,633.87) | 8,541.21 (7,627.92) | 278.52 (264.87) | 77.61 (217.21)   |
| None <sup>a</sup>          | -                          | -                 | -                   | -               | -                |
| <b>Unmet Need</b>          |                            |                   |                     |                 |                  |
| Yes                        | <b>8,186.00*(3,402.78)</b> | 1,270.29 (743.21) | 6,915.71 (3,479.80) | 142.69 (125.14) | -111.12 (101.28) |
| No <sup>a</sup>            | -                          | -                 | -                   | -               | -                |
| * p < 0.05                 |                            |                   |                     |                 |                  |

<sup>a</sup> Denotes referent group.

<sup>b</sup> Black and Other races were collapsed together to form “Non-white” race.

<sup>c</sup> “Two or more” and “one” healthcare coverages were collapsed together to form “One or more” healthcare coverage.

**Table 4. Adjusted Estimated Parameters from the Annual Healthcare Expenditures Model (in US Dollars, \$2005)**

Linear Regression using Ordinary Least Squares (OLS) results

| <i>Annual Healthcare Expenditures</i> |                                      |                                      |                                      |   |                                      |
|---------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|---|--------------------------------------|
|                                       | <b>Total</b>                         | <b>Inpatient<br/>Hospital</b>        | <b>Provider<br/>Encounters</b>       | <b>Over-the-counter<br/>Medications</b> | <b>Other<br/>Health Costs</b>        |
| <b>Independent Variables</b>          | <b>Estimate<br/>(Standard Error)</b> | <b>Estimate<br/>(Standard Error)</b> | <b>Estimate<br/>(Standard Error)</b> | <b>Estimate<br/>(Standard Error)</b>    | <b>Estimate<br/>(Standard Error)</b> |
| <b>CFS Registry</b>                   |                                      |                                      |                                      |   |                                      |
| Provider-referred                     | 2,462.15 (5,131.11)                  | -57.76 (1,198.11)                    | 2,519.91(5,231.18)                   | -36.25 (188.89)                         | -154.69 (144.45)                     |
| Self-referred <sup>a</sup>            | -                                    | -                                    | -                                    | -                                       | -                                    |
| <b>Sex</b>                            |                                      |                                      |                                      |   |                                      |
| Female                                | 6,327.61 (6,218.70)                  | 923.66 (1,452.06)                    | 5,403.95(6,339.98)                   | -198.90 (228.93)                        | <b>-382.00* (175.07)</b>             |
| Male <sup>a</sup>                     | -                                    | -                                    | -                                    | -                                       | -                                    |
| <b>Race</b>                           |                                      |                                      |                                      |   |                                      |
| White                                 | -1,789.84 (7,062.40)                 | 110.76 (1,649.06)                    | -1,900.60(7200.14)                   | 95.12 (259.99)                          | 228.75 (198.82)                      |
| Non-white <sup>a b</sup>              | -                                    | -                                    | -                                    | -                                       | -                                    |
| <b>Age groups</b>                     |                                      |                                      |                                      |   |                                      |
| 18-39                                 | 425.17 (4,163.24)                    | -599.41 (972.11)                     | 1,024.58(4,244.43)                   | 138.93 (153.26)                         | -33.79 (117.20)                      |
| 40-59 <sup>a</sup>                    | -                                    | -                                    | -                                    | -                                       | -                                    |
| 60-69                                 | 4,285.12 (6,398.46)                  | 1,140.94(1,494.03)                   | 3,144.18(6,523.25)                   | 253.34 (235.55)                         | -105.79 (180.13)                     |
| <b>Marital status</b>                 |                                      |                                      |                                      |   |                                      |
| Married                               | 3,007.90 (4,828.17)                  | -148.05 (1,127.37)                   | 3,155.96(4,922.33)                   | 169.81 (177.74)                         | 129.08 (135.92)                      |
| Unmarried <sup>a</sup>                | -                                    | -                                    | -                                    | -                                       | -                                    |
| <b>Education</b>                      |                                      |                                      |                                      |   |                                      |
| College degree or higher              | <b>8,932.15* (3,947.05)</b>          | -639.65 (921.63)                     | <b>9,571.80*(4,024.03)</b>           | 51.56 (145.30)                          | 119.12 (111.12)                      |
| Less than college degree <sup>a</sup> | -                                    | -                                    | -                                    | -                                       | -                                    |
| <b>Working status</b>                 |                                      |                                      |                                      |   |                                      |
| Full or part-time <sup>a</sup>        | -                                    | -                                    | -                                    | -                                       | -                                    |
| Not working                           | 857.38 (3,846.40)                    | 233.28 (898.13)                      | 624.10 (3,921.42)                    | 142.80 (141.60)                         | 126.54 (108.28)                      |

|                            |                             |                   |                            |                 |                  |
|----------------------------|-----------------------------|-------------------|----------------------------|-----------------|------------------|
| <b>Healthcare coverage</b> |                             |                   |                            |                 |                  |
| One or more <sup>c</sup>   | -876.41 (8,953.45)          | 52.70 (2,090.62)  | -929.11 (9,128.07)         | -23.98 (329.60) | -122.11 (252.05) |
| None <sup>a</sup>          | -                           | -                 | -                          | -               | -                |
| <b>Unmet Need</b>          |                             |                   |                            |                 |                  |
| Yes                        | <b>9,720.54* (3,890.10)</b> | 1,265.66 (908.33) | <b>8,454.88*(3,965.97)</b> | 99.80 (143.21)  | -123.94 (109.51) |
| No <sup>a</sup>            | -                           | -                 | -                          | -               | -                |
| * p < 0.05                 |                             |                   |                            |                 |                  |

<sup>a</sup> Denotes referent group.

<sup>b</sup> Black and Other races were collapsed together to form “Non-white” race.

<sup>c</sup> “Two or more” and “one” healthcare coverages were collapsed together to form “One or more” healthcare coverage.

## ANNOTATED SAS CODE

### Proc contents procedure for Table 1

The SAS System 12:58 Monday, April 9, 2012 1

#### The CONTENTS Procedure

|                     |                                    |                      |     |
|---------------------|------------------------------------|----------------------|-----|
| Data Set Name       | PUFLIB.TABLE1_WITH_HU_6            | Observations         | 35  |
| Member Type         | DATA                               | Variables            | 27  |
| Engine              | V8                                 | Indexes              | 0   |
| Created             | Monday, April 09, 2012 12:59:52 PM | Observation Length   | 200 |
| Last Modified       | Monday, April 09, 2012 12:59:52 PM | Deleted Observations | 0   |
| Protection          |                                    | Compressed           | NO  |
| Data Set Type       |                                    | Sorted               | NO  |
| Label               |                                    |                      |     |
| Data Representation | WINDOWS_32                         |                      |     |
| Encoding            | wlatin1 Western (Windows)          |                      |     |

#### Engine/Host Dependent Information

|                            |                                   |
|----------------------------|-----------------------------------|
| Data Set Page Size         | 16384                             |
| Number of Data Set Pages   | 1                                 |
| First Data Page            | 1                                 |
| Max Obs per Page           | 81                                |
| Obs in First Data Page     | 35                                |
| Number of Data Set Repairs | 0                                 |
| Filename                   | c:\meps\table1_with_hu_6.sas7bdat |
| Release Created            | 9.0202M3                          |
| Host Created               | XP_PRO                            |

#### Alphabetic List of Variables and Attributes

| # Variable     | Type | Len | Format | Informat | Label         |
|----------------|------|-----|--------|----------|---------------|
| 25 HU_6        | Num  | 8   |        |          | HU_6          |
| 17 SUPPORT_    | Char | 3   | \$3.   | \$3.     | SUPPORT_GROUP |
| GROUP          |      |     |        |          |               |
| 1 abtid        | Num  | 8   |        |          | abtid         |
| 5 age          | Num  | 8   |        |          | age           |
| 4 age_cate     | Num  | 8   |        |          | age_cate      |
| 19 annual_     | Num  | 8   |        |          |               |
| earnings       |      |     |        |          |               |
| 7 education    | Num  | 8   |        |          | education     |
| 8 ei_6         | Num  | 8   |        |          | ei_6          |
| 9 ei_16        | Num  | 8   |        |          | ei_16         |
| 10 ei_17       | Num  | 8   |        |          | ei_17         |
| 11 ei_19       | Num  | 8   |        |          | ei_19         |
| 12 ei_51_1     | Num  | 8   |        |          | ei_51_1       |
| 13 ei_51_2     | Num  | 8   |        |          | ei_51_2       |
| 14 ei_51_3     | Char | 1   | \$1.   | \$1.     | ei_51_3       |
| 15 ei_51_4     | Num  | 8   |        |          | ei_51_4       |
| 16 ei_51_5     | Char | 1   | \$1.   | \$1.     | ei_51_5       |
| 27 hu_met_need | Num  | 8   |        |          |               |
| 26 hu_unmet_   | Num  | 8   |        |          |               |
| need           |      |     |        |          |               |

The SAS System 12:58 Monday, April 9, 2012 2

#### The CONTENTS Procedure

#### Alphabetic List of Variables and Attributes

| #  | Variable     | Type | Len | Format | Informat | Label   |
|----|--------------|------|-----|--------|----------|---|
| 18 | income_      | Num  | 8   |        |          | Earnings from work during the past 4 weeks                                      |
|    | fourweek     |      |     |        |          |   |
| 20 | insure_new   | Num  | 8   |        |          | Sums non-missing health insurance vars to evaluate multiple healthcare coverage |
| 24 | insure_new_i | Num  | 8   |        |          |   |
| 6  | marital_     | Num  | 8   |        |          | marital_status  |
|    | status       |      |     |        |          |   |
| 3  | race         | Char | 5   | \$5.   | \$5.     | race  |
| 2  | sex          | Num  | 8   |        |          | sex   |
| 22 | support_     | Num  | 8   |        |          |   |
|    | group_i      |      |     |        |          |   |
| 21 | work_status  | Char | 9   |        |          |   |
| 23 | work_status_ | Num  | 8   |        |          |   |
|    | i            |      |     |        |          |   |

```

*****;
* Program: Table 1 variables extraction.sas          *;
* Date: 06.20.11                                  *;
* Programmer: Jeanie Lo                            *;
*                                                  *;
* Purpose: This program extracts the necessary    *;
* variables to create Table 1.                    *;
*****;

libname analysis '\\cdc.gov\private\M319\vkd2\SAS';
ods graphics on;

*** Creates a format for health insurance variables;
proc format;
    value insurancef 0    = "None"
                  1     = "One"
                  2-3   = "Two or More";
run;

*** Creates a permanent dataset called Table1_extraction in the
analysis library to begin variable extraction for Table 1;
data analysis.Table1_variables;
    set analysis.Table1_extraction;

*To calculate the 4 week earnings;
    if      ei_16 eq 1 then income_fourweek = ei_17*ei_19;
    else if ei_16 eq 2 then income_fourweek = 0;

    label income_fourweek = "Earnings from work during the past
    4 weeks";

*To calculate the annual earnings;
    annual_earnings = income_fourweek*13;

*To sum the health insurance variables to find out if subjects
have multiple insurances;
    insure_new = sum(ei_51_1, ei_51_2, ei_51_3, ei_51_4);

    label insure_new = "Sums non-missing health insurance vars
    to evaluate multiple healthcare coverage";

*To categorize the people without any health insurance;
    if ei_51_5 eq 1 then insure_new = 0;

*Creates a new variable named work_status that categorizes work
levels;
    if      ei_6 eq 1 then work_status = "Full-time";
    else if ei_6 eq 2 then work_status = "Part-time";
    else if ei_6 eq 3 then work_status = "Not currently
working";
run;

```

```
*** Prints data for sex;
*** Results: 4 males, 31 females;
proc freq data = analysis.Table1_variables;
    tables sex / list missing;
run;

*** Prints data for race;
*** Results: 31 White, 3 Black, and 1 Other;
proc freq data = analysis.Table1_variables;
    tables race / list missing;
run;

*** Prints data for age group;
*** Results: 18 to 29 = 2
           30 to 39 = 10
           40 to 49 = 5
           50 to 59 = 10
           60 to 69 = 8;
*** Note: Youngest subject = 24, oldest subject = 69;
proc freq data = analysis.Table1_variables;
    tables age_cate / list missing;
run;

*** Prints mean and standard deviation for age;
*** Results: Mean = 48.4, Std. = 12.39;
proc means data = analysis.Table1_variables;
    var age;
run;

*** Prints data for marital status;
*** Results: 26 Current married, 9 Not currently married;
proc freq data = analysis.Table1_variables;
    tables marital_status / list missing;
run;

*** Prints data for education;
*** Results: 22 College degree or higher, 13 Less than a college
degree;
proc freq data = analysis.Table1_variables;
    tables education / list missing;
run;

*** Prints data for working status;
*** Results: 10 Full-time, 5 Part-time, and 20 Not currently
working;
proc freq data = analysis.Table1_variables;
    tables work_status / list missing;
run;
```

```

*** Prints data for health insurance coverage;
*** Results: 4 Two or more, 29 One, and 2 None;
proc print data = analysis.Table1_variables;
    format insure_new insurancecf.;
    title 'Healthcare coverage';
run;

*** Prints data for % who worked at all in last 4 weeks (ei_16);
*** Results: 16/35 ~ 46%;
proc freq data = analysis.Table1_variables;
    tables ei_16 / list missing;
    where ei_16 eq 1;
run;

*** Prints data for annual earnings;
*** Results: $12,821;
proc means data = analysis.Table1_variables;
    var annual_earnings;
run;

*****
* START OF STRATIFICATION BY PROVIDER-REFERRED *
* * * * *
*****

*** Create a new permanent dataset that stratifies Registry CFS
subjects by referral type (provider vs. self);
data analysis.Table1_strat;
    set analysis.Table1_variables;

*Added support_group variable into Table1_variables spreadsheet
and re-imported data;
*support_group = "Yes" -> Self-referred;
*support_group = "No" -> Provider-referred;

*To create an indicator variable for support_group;
*support_group_i = 1 --> Yes (self-referred), support_group_i = 0
-> No (Provider-referred);
    if support_group = " " then support_group_i = 0;
    else if support_group = "Yes" then support_group_i = 1;

*To create an indicator variable for work_status;
*work_status_i = 0 --> Not currently working, work_status_i = 1 -
> Full-time, work_status_i = 2 -> Part-time;
    if work_status = "Not curre" then work_status_i = 0;
    else if work_status = "Full-time" then work_status_i = 1;
    else if work_status = "Part-time" then work_status_i = 2;

*To create an indicator variable for insure_new (multiple
healthcare coverage);

```



```

*insure_new_i = 0 -> None, insure_new_i = 1 -> One, insure_new_i
= 2 -> Two or more;
    if insure_new =      0    then insure_new_i = 0;
    else if insure_new = 1    then insure_new_i = 1;
    else if insure_new = 2 or insure_new = 3 then insure_new_i
    = 2;

* To create an indicator variable for ei_16 (% who worked at all
in last 4 weeks);
* ei_16_indic = 0 --> No, ei_16_indic = 1 --> Yes, where ei_16 =
1 --> "Yes" and ei_16 = 2 --> "No";
    if      ei_16 = 2 then ei_16_indic = 0;
    else if ei_16 = 1 then ei_16_indic = 1;
run;

*** Prints data for sex (Provider-referred);
*** Results: 1 male, 11 females;
proc freq data = analysis.Table1_strat;
    tables sex / list missing;
    where support_group = " ";
run;

*** Prints data for race (Provider-referred);
*** Results:  22 White, 1 Black, and 0 Other;
proc freq data = analysis.Table1_strat;
    tables race / list missing;
    where support_group = " ";
run;

*** Prints data for age group (Provider-referred);
*** Results: 18 to 29 = 2
                30 to 39 = 8
                40 to 49 = 4
                50 to 59 = 7
                60 to 69 = 2;
*** Note: Youngest subject = 24, oldest subject = 69;

proc freq data = analysis.Table1_strat;
    tables age_cate / list missing;
    where support_group = " ";
run;

*** Prints mean and standard deviation for age (Provider-
referred);
*** Results: Mean = 45.3, Std. = 11.83;
proc means data = analysis.Table1_strat;
    var age;
    where support_group = " ";
run;

```

```
*** Prints data for marital status (Provider-referred);
*** Results:  Current married,  Not currently married;
proc freq data = analysis.Table1_strat;
    tables marital_status / list missing;
    where support_group = " ";
run;

*** Prints data for education (Provider-referred);
*** Results:  10 College degree or higher, 13 Less than a college
degree;
proc freq data = analysis.Table1_strat;
    tables education / list missing;
    where support_group = " ";
run;

*** Prints data for working status (Provider-referred);
*** Results:  8 Full-time, 3 Part-time, and 12 Not currently
working;
proc freq data = analysis.Table1_strat;
    tables work_status / list missing;
    where support_group = " ";
run;

*** Prints data for health insurance coverage (Provider-
referred);
*** Results:  Two or more,  One, and  None;
proc freq data = analysis.Table1_strat;
    format insure_new insurancecf.;
    where support_group = " ";
run;

*** Prints data for % who worked at all in last 4 weeks (ei_16),
(Provider-referred);
*** Results:  12/35 ~ 34%;
proc freq data = analysis.Table1_strat;
    tables ei_16 / list missing;
    where ei_16 eq 1;
    where support_group = " ";
run;

*** Prints data for annual earnings (Provider-referred);
*** Results:  $15,572.87;
proc means data = analysis.Table1_strat;
    var annual_earnings;
    where support_group = " ";
run;
```

```

*****
*
* START OF STRATIFICATION BY SELF-REFERRED
*
*****

*** Prints data for sex (Self-referred);
*** Results: 1 male, 11 females;
proc freq data = analysis.Table1_strat;
    tables sex / list missing;
    where support_group = "Yes";
run;

*** Prints data for race (Self-referred);
*** Results: 9 White, 2 Black, and 1 Other;
proc freq data = analysis.Table1_strat;
    tables race / list missing;
    where support_group = "Yes";
run;

*** Prints data for age group (Self-referred);
*** Results: 18 to 29 = 0
              30 to 39 = 2
              40 to 49 = 1
              50 to 59 = 3
              60 to 69 = 6;
*** Note: Youngest subject = 30, oldest subject = 68;
proc freq data = analysis.Table1_strat;
    tables age_cate / list missing;
    where support_group = "Yes";
run;

*** Prints mean and standard deviation for age (Self-referred);
*** Results: Mean = 54.2, Std. = 11.78;
proc means data = analysis.Table1_strat;
    var age;
    where support_group = "Yes";
run;

*** Prints data for marital status (Self-referred);
*** Results: 9 Currently married, 3 Not currently married;
proc freq data = analysis.Table1_strat;
    tables marital_status / list missing;
    where support_group = "Yes";
run;

```

```

*** Prints data for education (Self-referred);
*** Results: 3 College degree or higher, 9 Less than a college
degree;

```

```

proc freq data = analysis.Table1_strat;
    tables education / list missing;
    where support_group = "Yes";
run;

```

```

*** Prints data for working status (Self-referred);
*** Results: 2 Full-time, 2 Part-time, and 8 Not currently
working;

```

```

proc freq data = analysis.Table1_strat;
    tables work_status / list missing;
    where support_group = "Yes";
run;

```

```

*** Prints data for health insurance coverage (Self-referred);
*** Results: 1 Two or more, 10 One, and 1 None;

```

```

proc freq data = analysis.Table1_strat;
    format insure_new insurancef.;
    where support_group = "Yes";
run;

```

```

*** Prints data for % who worked at all in last 4 weeks (ei_16),
(Self-referred);

```

```

*** Results: 4/35 ~ 11%;
proc freq data = analysis.Table1_strat;
    tables ei_16 / list missing;
    where ei_16 eq 1;
    where support_group = "Yes";
run;

```

```

*** Prints data for annual earnings (Self-referred);
*** Results: $7,545.42;

```

```

proc means data = analysis.Table1_strat;
    var annual_earnings;
    where support_group = "Yes";
run;

```

```

*****
* P-VALUES
*
*****

```

```

*** To determine p-value for sex;
*** Since 50% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 1.00;
proc freq data = analysis.Table1_strat;
    tables support_group_i*sex / expected chisq;
run;

*** To determine p-value for race;
*** Since 67% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 0.11;
proc freq data = analysis.Table1_strat;
    tables support_group_i*race / expected chisq fisher;
run;

*** To determine p-value for age groups;
*** Since 70% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 0.11;
proc freq data = analysis.Table1_strat;
    tables support_group_i*age_cate / expected chisq fisher;
run;

*** To determine p-value for age;
*** Since the equality of variances test shows a p-value of 1.00,
we cannot reject the null hypothesis that the mean differences
between the two groups are equal. Thus, we will be using the
Pooled (equal variances) method.;
*** Results: p = 0.04;
proc ttest data = analysis.Table1_strat;
    class support_group_i;
    var age;
run;

*** To determine p-value for marital status;
*** Since 25% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 1.00;
proc freq data = analysis.Table1_strat;
    tables support_group_i*marital_status / expected chisq
fisher;
run;

*** To determine p-value for education;
*** Since 25% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 0.46;
proc freq data = analysis.Table1_strat;
    tables support_group_i*education / expected chisq fisher;
run;

```

```
*** To determine p-value for working status;
*** Since 50% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 0.60;
proc freq data = analysis.Table1_strat;
    tables support_group_i*work_status_i / expected chisq
fisher;
run;

*** To determine p-value for multiple healthcare coverage;
*** Since 67% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 1.00;
proc freq data = analysis.Table1_strat;
    tables support_group_i*insure_new_i / expected chisq
fisher;
run;

*** To determine p-value for % who worked at all in last 4 weeks;
*** Results: p = 0.29 using chisq test;
proc freq data = analysis.Table1_strat;
    tables support_group_i*ei_16_indic / expected chisq fisher;
run;

*** To determine p-value for employment/earnings;
*** Since the p-value for the Equality of Variances test is 0.31,
we fail to reject the null hypothesis that the variances are
equal;
*** Thus, we go with the Pooled (equal variances) method;
*** Results: p = 0.25 ;
proc ttest data = analysis.Table1_strat;
    class support_group_i;
    var annual_earnings;
run;
```

## Proc contents procedure for Table 2

The SAS System 12:58 Monday, April 9, 2012 11

### The CONTENTS Procedure

|                     |                                     |                      |     |
|---------------------|-------------------------------------|----------------------|-----|
| Data Set Name       | PUFLIB.TABLE2_HOSPITALIZATION_COSTS | Observations         | 35  |
| Member Type         | DATA                                | Variables            | 88  |
| Engine              | V8                                  | Indexes              | 0   |
| Created             | Monday, March 12, 2012 10:21:20 AM  | Observation Length   | 952 |
| Last Modified       | Monday, March 12, 2012 10:21:20 AM  | Deleted Observations | 0   |
| Protection          |                                     | Compressed           | NO  |
| Data Set Type       |                                     | Sorted               | YES |
| Label               |                                     |                      |     |
| Data Representation | WINDOWS_32                          |                      |     |
| Encoding            | wlatin1 Western (Windows)           |                      |     |

### Engine/Host Dependent Information

|                            |   |
|----------------------------|---|
| Data Set Page Size         | 16384   |
| Number of Data Set Pages   | 3   |
| First Data Page            | 1   |
| Max Obs per Page           | 17  |
| Obs in First Data Page     | 6   |
| Number of Data Set Repairs | 0   |
| Filename                   | c:\meps\table2_hospitalization_costs.sas7bdat |
| Release Created            | 9.0202M3                                      |
| Host Created               | XP_PRO  |

### Alphabetic List of Variables and Attributes

| #  | Variable | Type | Len | Format | Informat | Label    |
|----|----------|------|-----|--------|----------|----------|
| 1  | abtid    | Num  | 8   |        |          | abtid    |
| 61 | ei_55    | Num  | 8   |        |          | ei_55    |
| 33 | ei_53a1  | Char | 28  | \$28.  | \$28.    | ei_53a1  |
| 34 | ei_53a2  | Char | 26  | \$26.  | \$26.    | ei_53a2  |
| 35 | ei_53a3  | Char | 38  | \$38.  | \$38.    | ei_53a3  |
| 36 | ei_53a4  | Char | 24  | \$24.  | \$24.    | ei_53a4  |
| 37 | ei_53a5  | Char | 27  | \$27.  | \$27.    | ei_53a5  |
| 38 | ei_53a6  | Char | 28  | \$28.  | \$28.    | ei_53a6  |
| 39 | ei_53a7  | Char | 28  | \$28.  | \$28.    | ei_53a7  |
| 40 | ei_53a8  | Char | 36  | \$36.  | \$36.    | ei_53a8  |
| 41 | ei_53a9  | Char | 27  | \$27.  | \$27.    | ei_53a9  |
| 42 | ei_53a10 | Char | 10  | \$10.  | \$10.    | ei_53a10 |
| 43 | ei_53a11 | Char | 20  | \$20.  | \$20.    | ei_53a11 |
| 44 | ei_53a12 | Char | 12  | \$12.  | \$12.    | ei_53a12 |
| 45 | ei_53a13 | Char | 17  | \$17.  | \$17.    | ei_53a13 |
| 46 | ei_53a14 | Char | 25  | \$25.  | \$25.    | ei_53a14 |
| 47 | ei_53b1  | Num  | 8   |        |          | ei_53b1  |
| 48 | ei_53b2  | Num  | 8   |        |          | ei_53b2  |
| 49 | ei_53b3  | Num  | 8   |        |          | ei_53b3  |
| 50 | ei_53b4  | Num  | 8   |        |          | ei_53b4  |
| 51 | ei_53b5  | Num  | 8   |        |          | ei_53b5  |
| 52 | ei_53b6  | Num  | 8   |        |          | ei_53b6  |

The SAS System 12:58 Monday, April 9, 2012 12

### The CONTENTS Procedure

### Alphabetic List of Variables and Attributes

| #  | Variable | Type | Len | Format | Informat | Label   |
|----|----------|------|-----|--------|----------|---------|
| 53 | ei_53b7  | Num  | 8   |        |          | ei_53b7 |
| 54 | ei_53b8  | Num  | 8   |        |          | ei_53b8 |
| 55 | ei_53b9  | Num  | 8   |        |          | ei_53b9 |

|    |               |      |    |       |       |              |
|----|---------------|------|----|-------|-------|--------------|
| 56 | ei_53b10      | Num  | 8  |       |       | ei_53b10     |
| 57 | ei_53b11      | Num  | 8  |       |       | ei_53b11     |
| 58 | ei_53b12      | Char | 2  | \$2.  | \$2.  | ei_53b12     |
| 59 | ei_53b13      | Char | 2  | \$2.  | \$2.  | ei_53b13     |
| 60 | ei_53b14      | Char | 2  | \$2.  | \$2.  | ei_53b14     |
| 62 | ei_57a1       | Char | 15 | \$15. | \$15. | ei_57a1      |
| 63 | ei_57a2       | Char | 8  | \$8.  | \$8.  | ei_57a2      |
| 64 | ei_57b1       | Num  | 8  |       |       | ei_57b1      |
| 65 | ei_57b2       | Num  | 8  |       |       | ei_57b2      |
| 69 | hospcost_ooop | Num  | 8  |       |       |              |
| 70 | hospcosttot   | Num  | 8  |       |       |              |
| 66 | hospdaysperpt | Num  | 8  |       |       |              |
| 67 | hospvisperpt  | Num  | 8  |       |       |              |
| 5  | hu_13a1       | Char | 35 | \$35. | \$35. | hu_13a1      |
| 6  | hu_13a2       | Char | 1  | \$1.  | \$1.  | hu_13a2      |
| 7  | hu_13a3       | Char | 1  | \$1.  | \$1.  | hu_13a3      |
| 8  | hu_13b1       | Char | 1  | \$1.  | \$1.  | hu_13b1      |
| 9  | hu_13c1       | Char | 1  | \$1.  | \$1.  | hu_13c1      |
| 10 | hu_13c2       | Char | 1  | \$1.  | \$1.  | hu_13c2      |
| 11 | hu_13c3       | Char | 1  | \$1.  | \$1.  | hu_13c3      |
| 12 | hu_14a1       | Num  | 8  |       |       | hu_14a1      |
| 13 | hu_14a2       | Num  | 8  |       |       | hu_14a2      |
| 14 | hu_14a3       | Num  | 8  |       |       | hu_14a3      |
| 15 | hu_14a4       | Num  | 8  |       |       | hu_14a4      |
| 17 | hu_14a5       | Num  | 8  |       |       | hu_14a5      |
| 19 | hu_14a6       | Num  | 8  |       |       | hu_14a6      |
| 21 | hu_14a7       | Char | 1  | \$1.  | \$1.  | hu_14a7      |
| 23 | hu_14a8       | Char | 1  | \$1.  | \$1.  | hu_14a8      |
| 16 | hu_14a4sp     | Char | 31 | \$31. | \$31. | hu_14a4sp    |
| 18 | hu_14a5sp     | Char | 26 | \$26. | \$26. | hu_14a5sp    |
| 20 | hu_14a6sp     | Char | 31 | \$31. | \$31. | hu_14a6sp    |
| 22 | hu_14a7sp     | Char | 12 | \$12. | \$12. | hu_14a7sp    |
| 24 | hu_14a8sp     | Char | 7  | \$7.  | \$7.  | hu_14a8sp    |
| 25 | hu_14b1       | Num  | 8  |       |       | hu_14b1      |
| 26 | hu_14b2       | Num  | 8  |       |       | hu_14b2      |
| 27 | hu_14b3       | Num  | 8  |       |       | hu_14b3      |
| 28 | hu_14b4       | Num  | 8  |       |       | hu_14b4      |
| 29 | hu_14b5       | Num  | 8  |       |       | hu_14b5      |
| 30 | hu_14b6       | Num  | 8  |       |       | hu_14b6      |
| 31 | hu_14b7       | Char | 1  | \$1.  | \$1.  | hu_14b7      |
| 32 | hu_14b8       | Char | 1  | \$1.  | \$1.  | hu_14b8      |
| 68 | i             | Num  | 8  |       |       |              |
| 4  | insure_new_i  | Num  | 8  |       |       | insure_new_i |
| 84 | j             | Num  | 8  |       |       |              |

The SAS System

12:58 Monday, April 9, 2012 13

## The CONTENTS Procedure

## Alphabetic List of Variables and Attributes

| #  | Variable         | Type | Len | Format | Informat | Label         |
|----|------------------|------|-----|--------|----------|---------------|
| 88 | otcandothtot     | Num  | 8   |        |          |               |
| 71 | otccosts_ooop    | Num  | 8   |        |          |               |
| 72 | otccoststot      | Num  | 8   |        |          |               |
| 73 | othhccosts_ooop  | Num  | 8   |        |          |               |
| 74 | othhccoststot    | Num  | 8   |        |          |               |
| 80 | provallcost_ooop | Num  | 8   |        |          |               |
| 81 | provallcosttot   | Num  | 8   |        |          |               |
| 78 | provaltcost_vis  | Num  | 8   |        |          |               |
| 79 | provdenccost_vis | Num  | 8   |        |          |               |
| 75 | provmdcost_vis   | Num  | 8   |        |          |               |
| 77 | provpsycost_vis  | Num  | 8   |        |          |               |
| 76 | provrncost_vis   | Num  | 8   |        |          |               |
| 82 | rxgencosts_ooop  | Num  | 8   |        |          |               |
| 83 | rxperpt          | Num  | 8   |        |          |               |
| 2  | support_group    | Char | 3   | \$3.   | \$3.     | support_group |



|    |                 |     |   |                 |
|----|-----------------|-----|---|-----------------|
| 3  | support_group_i | Num | 8 | support_group_i |
| 85 | totalcost       | Num | 8 |                 |
| 86 | totoopcost      | Num | 8 |                 |
| 87 | visityrly       | Num | 8 |                 |

## Sort Information

|               |       |
|---------------|-------|
| Sortedby      | abtid |
| Validated     | YES   |
| Character Set | ANSI  |

```

*****;
* Program: Merging MEPS.sas *;
* Date: 08.02.11 *;
* Programmer: Jeanie Lo *;
* *;
* Purpose: This program combines the 2005 Hospital *;
* Inpatient Stays File (HC 094D) with the 2005 Full *;
* Year Population Characteristics (HC-097) into one *;
* file. *;
*****;

libname puflib v8 'c:\meps';
libname in1 xport 'c:\meps\h94d.ssp';
libname in2 xport 'c:\meps\h97.ssp';
ods graphics on;

proc copy in = in1 out = puflib memtype = data;
run;

proc copy in = in2 out = puflib memtype = data;
run;

*** Creates a permanent dataset called h97_keep which only keeps
dupersid, inscov05, age05x, region05, and msa05 variables;
data puflib.h97_keep (keep = dupersid inscov05 age05x region05
msa05);
    set puflib.h97;
run;

*** Sort h94d by dupersid;
proc sort data = puflib.h94d;
    by dupersid;
run;

*** Sort h97_keep by dupersid;
proc sort data = puflib.h97_keep;
    by dupersid;
run;

*** Creates a permanent dataset called hospital_fees_revised in
the puflib library to construct hospital costs (revised version);
data puflib.hospital_fees_revised;

*Merge h94d and h97_keep by dupersid;
    merge puflib.h94d puflib.h97_keep;
    by dupersid;

```



```

*** To calculate mean total expenditures for non-metro and had an
operation;
*** This number will be multiplied with the mean hospdaysperpt
from the registry data and further stratified by provider type;
*** Results: $3,836.19;

```

```

proc means data = puflib.hospital_fees_revised;
    where msa05 eq 0 and anyoper eq 1;
    var xp;
run;

```

```

*****
* TABLE 2 - PROVIDER ENCOUNTERS MEAN COSTS
*
*****

```

```

data puflib.Table2_provider_costs;
    set puflib.h97;

```

```

*Remember 5 categories:
    1. ER
    2. Outpatient
    3. Office-based
    4. Dental
    5. Home health visits;

```

```

*Creates a new variable "total_ER_cost" to calculate sum of ER
facility and ER doctors' expenses costs;
    total_ER_cost = erfexp05 + erdexp05;

```

```

*Creates a new variable "total_outpatient_cost" to calculate sum
of OP facility expenses and OP provider expenses;
    total_outpatient_cost = opfexp05 + opdexp05;

```

```

*Creates a new variable "total_dental_cost" to calculate total
dental care visit expenses;
    total_dental_cost = dvtttch05;

```

```

*Creates a new variable "total_home_health_visit_cost" to
calculate total home health visit expenses;
    total_home_health_visit_cost = hhaexp05 + hhnexp05;

```

```

run;

```

```

*****;
* Program: Imputing Costs.sas *;
* Date: 08.10.11 *;
* Programmer: Jeanie Lo *;
* *;
* Purpose: This program calculates the mean costs for *;
* Table 2. *;
*****;

libname puflib v8 'c:\meps';
ods graphics on;

data puflib.Table2_hospitalization_costs (drop = f66 f67 f68 f69
f70 f71 f72 otccosttot othhcoststot);
    set puflib.Table2;

*Creates a new variable "hospdaysperpt" to calculate number of
inpatient hospital nights in past 1 year;
    hospdaysperpt = sum(hu_13c1, hu_13c2, hu_13c3);

*Sets hospdaysperpt to 0 if there are missing values;
    if hospdaysperpt eq . then hospdaysperpt = 0;

*Creates a new variable "hospvisperpt" to calculate number of
inpatient hospital visits in past 1 year and initialize it to 0;
    hospvisperpt = 0;

*Creates an array named "hospital_visits" to count the non-
missing hospital visits;
    array hospital_visits {3} hu_13a1 hu_13a2 hu_13a3;

*Use a do loop to read through the array 3 times (once for each
variable),and use a variable called "i" to keep track of how many
times we loop through;
    do i = 1 to 3;

        if (hospital_visits {i} ne '') then
            hospvisperpt=hospvisperpt+1;
        if (hospital_visits {i} eq '') then
            hospvisperpt=hospvisperpt+0;
        end;
    end;

```

```
*Creates a new variable "hospcost_oop" to calculate the OOP cost
of inpatient stays;
*Note: Only 2 subjects were hospitalized. Both didn't have
surgical procedures;
*Note: The $39 figure is from p. 14 of the economic analysis
binder using Non-MSA, Insured, Non-surgical, and OOP;
    hospcost_oop = hospdaysperpt * 39;

*Sets hospcost_oop to 0 if there are missing values;
    if hospcost_oop eq . then hospcost_oop = 0;

*Creates a new variable "hospcosttot" to calculate the total cost
of inpatient stays;
*Note: The $1,976 figure is from p. 14 of the economic analysis
binder using Non-MSA, Insured, Non-surgical, and Total
Expenditure;
    hospcosttot = hospdaysperpt * 1976;

*Sets hospcosttot to 0 if there are missing values;
    if hospcosttot eq . then hospcosttot = 0;

*Creates a new variable "otccosts_oop" to calculate the OOP cost
of OTC meds in past 1 year;
    otccosts_oop = ei_55 * 13;

*Sets ei_55 to 0 if there are missing values;
    if ei_55 eq . then ei_55 = 0;

*Sets otccosts_oop to 0 if there are missing values;
    if otccosts_oop eq . then otccosts_oop = 0;

*Creates a new variable "otccoststot" to calculate the total cost
of OTC meds in past 1 year;
    otccoststot = ei_55 * 13;

*Sets ei_55 to 0 if there are missing values;
    if ei_55 eq . then ei_55 = 0;

*Sets otccoststot to 0 if there are missing values;
    if otccoststot eq . then otccoststot = 0;

*Sets ei_57b1 and ei_57b2 to 0 if there are missing values;
    if ei_57b1 eq . then ei_57b1 = 0;
    if ei_57b2 eq . then ei_57b2 = 0;
```

\*Creates a new variable "othhccosts\_oop" to calculate OOP other health-related expenditures in past 1 year;

```
othhccosts_oop = ((sum(ei_57b1, ei_57b2)) * 13);
```

\*Sets othhccosts\_oop to 0 if there are missing values;

```
if othhccosts_oop eq . then othhccosts_oop = 0;
```

\*Creates a new variable "othhccoststot" to calculate total other health-related expenditures in past 1 year;

```
othhccoststot = ((sum(ei_57b1, ei_57b2)) * 13);
```

\*Sets othhccoststot to 0 if there are missing values;

```
if othhccoststot eq . then othhccoststot = 0;
```

\*Creates a new variable "provmdcost\_vis" to calculate the number of doctor visits per year;

```
if hu_14a1 eq 1 then provmdcost_vis = hu_14b1 * 2;
else if hu_14a1 eq 2 then provmdcost_vis = 0;
```

\*Creates a new variable "provrncost\_vis" to calculate the number of RN or PA visits per year;

```
if hu_14a2 eq 1 then provrncost_vis = hu_14b2 * 2;
else if hu_14a2 eq 2 then provrncost_vis = 0;
```

\*Creates a new variable "provpsycost\_vis" to calculate the number of psychiatrist, psychologist, or counselor visits per year;

```
if hu_14a3 eq 1 then provpsycost_vis = hu_14b3 * 2;
else if hu_14a3 eq 2 then provpsycost_vis = 0;
```

\*Creates a new variable "provaltcost\_vis" to calculate the number of other healthcare professional (e.g. optometrist, chiropractor, massage therapist) visits per year;

```
if hu_14a4 eq 1 then provaltcost_vis = hu_14b4 * 2;
else if hu_14a4 eq 2 then provaltcost_vis = 0;
```

```
if hu_14a5 eq 1 then provaltcost_vis = (hu_14b5 *
2) + provaltcost_vis;
```

```
if hu_14a6 eq 1 then provaltcost_vis = (hu_14b6 *
2) + provaltcost_vis;
```

\*Creates a new variable "provdencost\_vis" to calculate the number of dental visits per year;

```
if hu_14a4sp eq "Dentist" or
   hu_14a5sp eq "Dentist" or
   hu_14a5sp eq "Dentist, exam & root canal" or
   hu_14a8sp eq "Dentist" then provdencost_vis =
  (sum(hu_14b4, hu_14b5, hu_14b8)) * 2;
```

\*Sets provdencost\_vis to 0 if there are missing values;

```
if provdencost_vis eq . then provdencost_vis = 0;
```

\*Creates a new variable "provallcost\_oop" to calculate annualized OOP costs for all health care provider visits;

\*Note: The two subjects who didn't have health insurance were 990000974 and 990001060. All others had at least 1 form of health insurance;

\*Note: The \$26 figure is taken from p. 15 of the economic analysis binder for Non-MSA, OOP Expenditures, and Insured;

\*Note: The \$88 figure is taken from p. 15 of the economic analysis binder for Non-MSA, OOP Expenditures, and Uninsured;

```
if abtid ne 990000974 or abtid ne 990001060 then
  provallcost_oop=
  (sum(provmdcost_vis, provrncost_vis, provpsycost_vis,
        provdencost_vis, provaltcost_vis)) * 26;
if abtid eq 990000974 or abtid eq 990001060 then
provallcost_oop =
  (sum(provmdcost_vis, provrncost_vis, provpsycost_vis,
        provdencost_vis, provaltcost_vis)) * 88;
```

\*Creates a new variable "provallcosttot" to calculate annualized total costs for all health care provider encounters;

\*Note: The two subjects who didn't have health insurance were 990000974 and 990001060. All others had at least 1 form of health insurance;

\*Note: The \$202 figure is taken from p. 15 of the economic analysis binder for Non-MSA, Total Expenditures, and Insured;

\*Note: The \$ 92 figure is taken from p. 15 of the economic analysis binder for Non-MSA, Total Expenditures, and Uninsured;

```
if abtid ne 990000974 or abtid ne 990001060 then
  provallcosttot =

  (sum(provmdcost_vis, provrncost_vis, provpsycost_vis,
        provdencost_vis, provaltcost_vis)) * 202;
if abtid eq 990000974 or abtid eq 990001060 then
  provallcosttot =

  (sum(provmdcost_vis, provrncost_vis, provpsycost_vis,
        provdencost_vis, provaltcost_vis)) * 92;
```



```

*Creates a new variable "totalcost" that calculates total
healthcare expenditures in past 1 year;
    totalcost = sum(hospcosttot, provallcosttot, otccosttot,
    othhcoststot);

*Creates a new variable "totoopcost" that calculates total OOP
expenditures in past 1 year;
    totoopcost = sum(hospcost_oop, provallcost_oop);

*Creates a new variable "visityrly" to calculate the number of
visits to any outpatient provider in past 1 year;
    if      hu_14a1 eq 1 then visityrly = hu_14b1 * 2;
    else if hu_14a1 eq 2 then visityrly = 0;

    if      hu_14a2 eq 1 then visityrly = (hu_14b2 * 2) +
visityrly;
    if      hu_14a3 eq 1 then visityrly = (hu_14b3 * 2) +
visityrly;
    if      hu_14a4 eq 1 then visityrly = (hu_14b4 * 2) +
visityrly;
    if      hu_14a5 eq 1 then visityrly = (hu_14b5 * 2) +
visityrly;
    if      hu_14a6 eq 1 then visityrly = (hu_14b6 * 2) +
visityrly;
    if      hu_14a7 eq 1 then visityrly = (hu_14b7 * 2) +
visityrly;
    if      hu_14a8 eq 1 then visityrly = (hu_14b8 * 2) +
visityrly;

*Creates a new variable "otcandothtot" to calculate the total OTC
meds and other healthcare costs in past 1 year;
    if otccoststot ne . and othhccoststot ne . then
otcandothtot = otccoststot + othhccoststot;
    if otccoststot eq . and othhccoststot ne . then
otcandothtot = othhccoststot + 0;
    if othhccoststot eq . and otccoststot ne . then
otcandothtot = otccoststot + 0;
    if othhccoststot eq . and otccoststot eq . then
otcandothtot = 0;

run;

*** To check if the hospital_visits array worked;
proc freq data = puflib.Table2_hospitalization_costs;
    tables hospvisperpt / list missing;
run;

```

```
*****
* TABLE 2 - MEAN ANNUAL HEALTHCARE EXPENDITURES *
* * * * *
*****
```

```
*** To determine means for inpatient hospitalization costs;
*** Results: Provider-referred OOP mean: $5.09, Self-referred OOP
mean: $19.50;
*** Results: Provider-referred total mean: $257.80, Self-referred
total mean: $988;
```

```
proc ttest data = puflib.Table2_hospitalization_costs;
    class support_group_i;
    var hospcost_oop hospcosttot;
run;
```

```
*** To determine means for ambulatory provider visits;
*** Results: Provider-referred OOP mean: $1,536, Self-referred
OOP mean: $1,118;
*** Results: Provider-referred total mean: $11,213.20, Self-
referred total mean: $7,899.70;
```

```
proc ttest data = puflib.Table2_hospitalization_costs;
    class support_group_i;
    var provallcost_oop provallcosttot;
run;
```

```
*** To determine means for OTC meds and other healthcare costs;
*** Results: Provider-referred total mean: $349.90, Self-referred
total mean: $498.30;
```

```
proc ttest data = puflib.Table2_hospitalization_costs;
    class support_group_i;
    var otcandothtot;
run;
```

```
*** To determine means for other health costs;
*** Results: Provider-referred total mean: $49.17, Self-referred
total mean: $119.20;
```

```
proc ttest data = puflib.Table2_hospitalization_costs;
    class support_group_i;
    var othhccoststot;
run;
```

```
*** To determine means for OTC meds;
*** Results: Provider-referred total mean: $300.70, Self-referred
total mean: $379.20;
```

```
proc ttest data = puflib.Table2_hospitalization_costs;
    class support_group_i;
    var otccoststot;
run;
```

```

*** 03.05.12 - Plots histogram of inpatient hospital expenses by
healthcare coverage groups to check for skewness;
proc univariate data = puflib.Table2_hospitalization_costs normal
plot;
    var hospcosttot;
    class insure_new_i;
    histogram hospcosttot / normal;
    inset mean std;
    title justify = right "Histogram of inpatient hospital
    expenses by healthcare coverage groups";
run;

*** 3.12.12 - To check the number of females with other health
costs;
*** Sort Table2_hospitalization costs by abtid;
proc sort data = puflib.Table2_hospitalization_costs;
    by abtid;
run;

*** Sort Table1_with_hu_6 by abtid;
proc sort data = puflib.Table1_with_hu_6;
    by abtid;
run;

*** Creates a permanent dataset called
checking_other_health_costs in the puflib library to check the #
of females with other health costs;
data puflib.checking_other_health_costs;

*Merge Table2_hospitalization costs and Table1_with_hu_6 by
abtid;
    merge puflib.Table2_hospitalization_costs
puflib.Table1_with_hu_6;
    by abtid;
run;

*** Results: Only 2 females had other health costs greater than
0, and 1 man had other health costs greater than 0.;
proc freq data = puflib.checking_other_health_costs;
    tables sex*othhccoststot / list missing;
run;

```

## Proc contents procedure for Tables 3-4

The SAS System 12:58 Monday, April 9, 2012 7

### The CONTENTS Procedure

|                     |                                    |                      |      |
|---------------------|------------------------------------|----------------------|------|
| Data Set Name       | PUFLIB.TABLE3_WITH_HU_6            | Observations         | 35   |
| Member Type         | DATA                               | Variables            | 138  |
| Engine              | V8                                 | Indexes              | 0    |
| Created             | Monday, April 09, 2012 12:18:58 PM | Observation Length   | 1336 |
| Last Modified       | Monday, April 09, 2012 12:18:58 PM | Deleted Observations | 0    |
| Protection          |                                    | Compressed           | NO   |
| Data Set Type       |                                    | Sorted               | NO   |
| Label               |                                    |                      |      |
| Data Representation | WINDOWS_32                         |                      |      |
| Encoding            | wlatin1 Western (Windows)          |                      |      |

### Engine/Host Dependent Information

|                            |                                    |
|----------------------------|------------------------------------|
| Data Set Page Size         | 16384                              |
| Number of Data Set Pages   | 4                                  |
| First Data Page            | 2                                  |
| Max Obs per Page           | 12                                 |
| Obs in First Data Page     | 12                                 |
| Number of Data Set Repairs | 0                                  |
| Filename                   | c:\meps\table3_with_hu_6.sas7bdatt |
| Release Created            | 9.0202M3                           |
| Host Created               | XP_PRO                             |

### Alphabetic List of Variables and Attributes

| # Variable              | Type | Len | Format | Informat | Label                                     |
|-------------------------|------|-----|--------|----------|---|
| 25 HU_6                 | Num  | 8   |        |          | HU_6                                      |
| 17 SUPPORT_GROUP        | Char | 3   | \$3.   | \$3.     | SUPPORT_GROUP                             |
| 1 abtid                 | Num  | 8   |        |          | abtid                                     |
| 5 age                   | Num  | 8   |        |          | age                                       |
| 119 age_18_29           | Num  | 8   |        |          | Indicator for age 18-29                   |
| 134 age_18_39           | Num  | 8   |        |          |   |
| 120 age_30_39           | Num  | 8   |        |          | Indicator for age 30-39                   |
| 121 age_40_49           | Num  | 8   |        |          | Indicator for age 40-49                   |
| 135 age_40_59           | Num  | 8   |        |          |   |
| 122 age_50_59           | Num  | 8   |        |          | Indicator for age 50-59                   |
| 123 age_60_69           | Num  | 8   |        |          | Indicator for age 60-69                   |
| 4 age_cate              | Num  | 8   |        |          | age_cate                                  |
| 19 annual_earnings      | Num  | 8   |        |          |   |
| 7 education             | Num  | 8   |        |          | education                                 |
| 126 education_college   | Num  | 8   |        |          | Indicator for college education           |
| 127 education_nocollege | Num  | 8   |        |          | Indicator for less than college education |
| 8 ei_6                  | Num  | 8   |        |          | ei_6                                      |
| 9 ei_16                 | Num  | 8   |        |          | ei_16                                     |
| 10 ei_17                | Num  | 8   |        |          | ei_17                                     |

The SAS System 12:58 Monday, April 9, 2012 8

### The CONTENTS Procedure

### Alphabetic List of Variables and Attributes

| # Variable | Type | Len | Format | Informat | Label   |
|------------|------|-----|--------|----------|---------|
| 11 ei_19   | Num  | 8   |        |          | ei_19   |
| 84 ei_55   | Num  | 8   |        |          | ei_55   |
| 12 ei_51_1 | Num  | 8   |        |          | ei_51_1 |

|                  |      |    |       |       |          |
|------------------|------|----|-------|-------|----------|
| 13 ei_51_2       | Num  | 8  |       |       | ei_51_2  |
| 14 ei_51_3       | Char | 1  | \$1.  | \$1.  | ei_51_3  |
| 15 ei_51_4       | Num  | 8  |       |       | ei_51_4  |
| 16 ei_51_5       | Char | 1  | \$1.  | \$1.  | ei_51_5  |
| 56 ei_53a1       | Char | 28 | \$28. | \$28. | ei_53a1  |
| 57 ei_53a2       | Char | 26 | \$26. | \$26. | ei_53a2  |
| 58 ei_53a3       | Char | 38 | \$38. | \$38. | ei_53a3  |
| 59 ei_53a4       | Char | 24 | \$24. | \$24. | ei_53a4  |
| 60 ei_53a5       | Char | 27 | \$27. | \$27. | ei_53a5  |
| 61 ei_53a6       | Char | 28 | \$28. | \$28. | ei_53a6  |
| 62 ei_53a7       | Char | 28 | \$28. | \$28. | ei_53a7  |
| 63 ei_53a8       | Char | 36 | \$36. | \$36. | ei_53a8  |
| 64 ei_53a9       | Char | 27 | \$27. | \$27. | ei_53a9  |
| 65 ei_53a10      | Char | 10 | \$10. | \$10. | ei_53a10 |
| 66 ei_53a11      | Char | 20 | \$20. | \$20. | ei_53a11 |
| 67 ei_53a12      | Char | 12 | \$12. | \$12. | ei_53a12 |
| 68 ei_53a13      | Char | 17 | \$17. | \$17. | ei_53a13 |
| 69 ei_53a14      | Char | 25 | \$25. | \$25. | ei_53a14 |
| 70 ei_53b1       | Num  | 8  |       |       | ei_53b1  |
| 71 ei_53b2       | Num  | 8  |       |       | ei_53b2  |
| 72 ei_53b3       | Num  | 8  |       |       | ei_53b3  |
| 73 ei_53b4       | Num  | 8  |       |       | ei_53b4  |
| 74 ei_53b5       | Num  | 8  |       |       | ei_53b5  |
| 75 ei_53b6       | Num  | 8  |       |       | ei_53b6  |
| 76 ei_53b7       | Num  | 8  |       |       | ei_53b7  |
| 77 ei_53b8       | Num  | 8  |       |       | ei_53b8  |
| 78 ei_53b9       | Num  | 8  |       |       | ei_53b9  |
| 79 ei_53b10      | Num  | 8  |       |       | ei_53b10 |
| 80 ei_53b11      | Num  | 8  |       |       | ei_53b11 |
| 81 ei_53b12      | Char | 2  | \$2.  | \$2.  | ei_53b12 |
| 82 ei_53b13      | Char | 2  | \$2.  | \$2.  | ei_53b13 |
| 83 ei_53b14      | Char | 2  | \$2.  | \$2.  | ei_53b14 |
| 85 ei_57a1       | Char | 15 | \$15. | \$15. | ei_57a1  |
| 86 ei_57a2       | Char | 8  | \$8.  | \$8.  | ei_57a2  |
| 87 ei_57b1       | Num  | 8  |       |       | ei_57b1  |
| 88 ei_57b2       | Num  | 8  |       |       | ei_57b2  |
| 92 hospcost_oop  | Num  | 8  |       |       |          |
| 93 hospcosttot   | Num  | 8  |       |       |          |
| 89 hospdaysperpt | Num  | 8  |       |       |          |
| 90 hospvisperpt  | Num  | 8  |       |       |          |
| 28 hu_13a1       | Char | 35 | \$35. | \$35. | hu_13a1  |
| 29 hu_13a2       | Char | 1  | \$1.  | \$1.  | hu_13a2  |
| 30 hu_13a3       | Char | 1  | \$1.  | \$1.  | hu_13a3  |
| 31 hu_13b1       | Char | 1  | \$1.  | \$1.  | hu_13b1  |

## The CONTENTS Procedure

## Alphabetic List of Variables and Attributes

| #   | Variable       | Type | Len | Format | Informat | Label   |
|-----|----------------|------|-----|--------|----------|---|
| 32  | hu_13c1        | Char | 1   | \$1.   | \$1.     | hu_13c1   |
| 33  | hu_13c2        | Char | 1   | \$1.   | \$1.     | hu_13c2   |
| 34  | hu_13c3        | Char | 1   | \$1.   | \$1.     | hu_13c3   |
| 35  | hu_14a1        | Num  | 8   |        |          | hu_14a1   |
| 36  | hu_14a2        | Num  | 8   |        |          | hu_14a2   |
| 37  | hu_14a3        | Num  | 8   |        |          | hu_14a3   |
| 38  | hu_14a4        | Num  | 8   |        |          | hu_14a4   |
| 40  | hu_14a5        | Num  | 8   |        |          | hu_14a5   |
| 42  | hu_14a6        | Num  | 8   |        |          | hu_14a6   |
| 44  | hu_14a7        | Char | 1   | \$1.   | \$1.     | hu_14a7   |
| 46  | hu_14a8        | Char | 1   | \$1.   | \$1.     | hu_14a8   |
| 39  | hu_14a4sp      | Char | 31  | \$31.  | \$31.    | hu_14a4sp   |
| 41  | hu_14a5sp      | Char | 26  | \$26.  | \$26.    | hu_14a5sp   |
| 43  | hu_14a6sp      | Char | 31  | \$31.  | \$31.    | hu_14a6sp   |
| 45  | hu_14a7sp      | Char | 12  | \$12.  | \$12.    | hu_14a7sp   |
| 47  | hu_14a8sp      | Char | 7   | \$7.   | \$7.     | hu_14a8sp   |
| 48  | hu_14b1        | Num  | 8   |        |          | hu_14b1   |
| 49  | hu_14b2        | Num  | 8   |        |          | hu_14b2   |
| 50  | hu_14b3        | Num  | 8   |        |          | hu_14b3   |
| 51  | hu_14b4        | Num  | 8   |        |          | hu_14b4   |
| 52  | hu_14b5        | Num  | 8   |        |          | hu_14b5   |
| 53  | hu_14b6        | Num  | 8   |        |          | hu_14b6   |
| 54  | hu_14b7        | Char | 1   | \$1.   | \$1.     | hu_14b7   |
| 55  | hu_14b8        | Char | 1   | \$1.   | \$1.     | hu_14b8   |
| 27  | hu_met_need    | Num  | 8   |        |          |   |
| 26  | hu_unmet_need  | Num  | 8   |        |          |   |
| 91  | i              | Num  | 8   |        |          |   |
| 18  | income_        | Num  | 8   |        |          | Earnings from work during the past 4 weeks                                      |
|     | fourweek       |      |     |        |          |   |
| 138 | insure         | Num  | 8   |        |          | Indicator for healthcare coverage   |
| 20  | insure_new     | Num  | 8   |        |          | Sums non-missing health insurance vars to evaluate multiple healthcare coverage |
| 24  | insure_new_i   | Num  | 8   |        |          | insure_new_i  |
| 133 | insure_none    | Num  | 8   |        |          | Indicator for having 0 healthcare coverage                                      |
| 132 | insure_one     | Num  | 8   |        |          | Indicator for having 1 healthcare coverage                                      |
| 131 | insure_two     | Num  | 8   |        |          | Indicator for having 2+ healthcare coverages                                    |
| 107 | j              | Num  | 8   |        |          |   |
| 124 | marital_       | Num  | 8   |        |          | Indicator for married status  |
|     | married        |      |     |        |          |   |
| 6   | marital_status | Num  | 8   |        |          | marital_status  |
| 125 | marital_       | Num  | 8   |        |          | Indicator for unmarried status  |
|     | unmarried      |      |     |        |          |   |
| 111 | otcandothtot   | Num  | 8   |        |          |   |
| 94  | otccosts_oop   | Num  | 8   |        |          |   |
| 95  | otccoststot    | Num  | 8   |        |          |   |
| 96  | othhccosts_oop | Num  | 8   |        |          |   |
| 97  | othhccoststot  | Num  | 8   |        |          |   |

The SAS System

12:58 Monday, April 9, 2012 10

## The CONTENTS Procedure

## Alphabetic List of Variables and Attributes

| #   | Variable       | Type | Len | Format | Informat | Label                               |
|-----|----------------|------|-----|--------|----------|-------------------------------------|
| 103 | provalcost_    | Num  | 8   |        |          |                                     |
|     | oop            |      |     |        |          |                                     |
| 104 | provalcosttot  | Num  | 8   |        |          |                                     |
| 101 | provalcost_    | Num  | 8   |        |          |                                     |
|     | vis            |      |     |        |          |                                     |
| 102 | provdencost_   | Num  | 8   |        |          |                                     |
|     | vis            |      |     |        |          |                                     |
| 98  | provmdcost_vis | Num  | 8   |        |          |                                     |
| 100 | provpsycost_   | Num  | 8   |        |          |                                     |
|     | vis            |      |     |        |          |                                     |
| 99  | provrncost_vis | Num  | 8   |        |          |                                     |
| 3   | race           | Char | 5   | \$5.   | \$5.     | race                                |
| 117 | race_black     | Num  | 8   |        |          | Indicator for black race            |
| 137 | race_nonwhite  | Num  | 8   |        |          | Indicator for non-white race        |
| 118 | race_other     | Num  | 8   |        |          | Indicator for other race            |
| 116 | race_white     | Num  | 8   |        |          | Indicator for white race            |
| 105 | rxgencosts_oop | Num  | 8   |        |          |                                     |
| 106 | rxperpt        | Num  | 8   |        |          |                                     |
| 2   | sex            | Num  | 8   |        |          | sex                                 |
| 114 | sex_female     | Num  | 8   |        |          | Indicator for female gender         |
| 115 | sex_male       | Num  | 8   |        |          | Indicator for male gender           |
| 22  | support_       | Num  | 8   |        |          | support_group_i                     |
|     | group_i        |      |     |        |          |                                     |
| 108 | totalcost      | Num  | 8   |        |          |                                     |
| 109 | totoopcost     | Num  | 8   |        |          |                                     |
| 112 | type_provider  | Num  | 8   |        |          | Indicator for provider referral     |
| 113 | type_self      | Num  | 8   |        |          | Indicator for self referral         |
| 110 | visityrly      | Num  | 8   |        |          |                                     |
| 21  | work_status    | Char | 9   |        |          |                                     |
| 23  | work_status_i  | Num  | 8   |        |          |                                     |
| 129 | working_full   | Num  | 8   |        |          | Indicator for working full-time     |
| 136 | working_       | Num  | 8   |        |          |                                     |
|     | fullandpart    |      |     |        |          |                                     |
| 130 | working_none   | Num  | 8   |        |          | Indicator for not currently working |
| 128 | working_part   | Num  | 8   |        |          | Indicator for working part-time     |

```

*****;
* Program: Table3.sas *;
* Date: 09.13.11 *;
* Programmer: Jeanie Lo *;
* *;
* Purpose: This program calculates the regression *;
* parameter estimates and standard errors for Table 3. *;
*****;

libname puflib v8 'c:\meps';
libname analysis '\\cdc.gov\private\M319\vkd2\SAS';
ods graphics on;

*** Creates a permanent dataset called Table1 in the puflib
library that is a copy of the Table 1 variables;
data puflib.Table1;
    set analysis.Table1_strat;
run;

*** Creates a permanent dataset called Table1_with_hu_6 in the
puflib library that adds the hu_6 variable;
data puflib.Table1_with_hu_6;
    merge puflib.Table1 (in = in1) analysis.hu (in = in2);
    by abtid;

    keep abtid sex race age_cate age marital_status education
    ei_6 ei_16 ei_17 ei_19 ei_51_1 ei_51_2 ei_51_3 ei_51_4
    ei_51_5 support_group income_fourweek annual_earnings
    insure_new work_status support_group_i work_status_i
    insure_new_i hu_6 hu_unmet_need hu_met_need;

    if in1 ne in2 then delete;

    if hu_6 eq 1 then hu_unmet_need = 1;
    else hu_unmet_need = 0;

    if hu_6 eq 2 then hu_met_need = 1;
    else hu_met_need = 0;
run;

*** Prints data for unmet need (Self-referred);
*** Results: 4 = had unmet need, 8 = needs were met;
proc freq data = puflib.Table1_with_hu_6;
    tables hu_6 / list missing;
    where support_group = "Yes";
run;

```



```

*** Prints data for unmet need (Provider-referred);
*** Results: 10 = had unmet need, 13 = needs were met;
proc freq data = puflib.Table1_with_hu_6;
    tables hu_6 / list missing;
    where support_group = " ";
run;

*** To determine p-value for unmet needs;
*** Since 25% of cells have expected cell counts < 5, will be
using Fisher's Test;
*** Results: p = 0.72;
proc freq data = puflib.Table1_with_hu_6;
    tables support_group_i*hu_6 / expected chisq;
run;

*** Sort the data prior to merging by abtid;
proc sort data = puflib.Table1;
    by abtid;
run;

proc sort data = puflib.Table2_hospitalization_costs;
    by abtid;
run;

*** Merges Table1 with Table2_hospitalization_costs by abtid to
create a new permanent dataset called Table3;
data puflib.Table3;
    merge puflib.Table1 puflib.Table2_hospitalization_costs;
    by abtid;

* Creates indicator variables for provider type;
    if support_group_i eq 0 then type_provider = 1;
    else type_provider = 0;

    if support_group_i eq 1 then type_self = 1;
    else type_self = 0;

    label type_provider = "Indicator for provider referral";
    label type_self      = "Indicator for self referral";

* Creates indicator variables for sex;
    if sex eq 2 then sex_female = 1;
    else sex_female = 0;

    if sex eq 1 then sex_male = 1;
    else sex_male = 0;

    label sex_female = "Indicator for female gender";
    label sex_male   = "Indicator for male gender";

```

```
* Creates indicator variables for race;
  if race eq 'White' then race_white = 1;
  else race_white = 0;

  if race eq 'Black' then race_black = 1;
  else race_black = 0;

  if race eq 'Other' then race_other = 1;
  else race_other = 0;

  label race_white = "Indicator for white race";
  label race_black = "Indicator for black race";
  label race_other = "Indicator for other race";

* Creates indicator variables for age;
  if age ge 18 and age le 29 then age_18_29 = 1;
  else age_18_29 = 0;

  if age ge 30 and age le 39 then age_30_39 = 1;
  else age_30_39 = 0;

  if age ge 40 and age le 49 then age_40_49 = 1;
  else age_40_49 = 0;

  if age ge 50 and age le 59 then age_50_59 = 1;
  else age_50_59 = 0;

  if age ge 60 and age le 69 then age_60_69 = 1;
  else age_60_69 = 0;

  label age_18_29 = "Indicator for age 18-29";
  label age_30_39 = "Indicator for age 30-39";
  label age_40_49 = "Indicator for age 40-49";
  label age_50_59 = "Indicator for age 50-59";
  label age_60_69 = "Indicator for age 60-69";

* Creates indicator variables for marital status;
  if marital_status eq 1 then marital_married = 1;
  else marital_married = 0;

  if marital_status eq 0 then marital_unmarried = 1;
  else marital_unmarried = 0;

  label marital_married = "Indicator for married status";
  label marital_unmarried = "Indicator for unmarried status";

* Creates indicator variables for education;
  if education eq 1 then education_college = 1;
  else education_college = 0;
```

```

if education eq 0 then education_nocollege = 1;
else education_nocollege = 0;

label education_college = "Indicator for college
education";
label education_nocollege = "Indicator for less than
college education";

* Creates indicator variables for working status;
if work_status_i eq 2 then working_part = 1;
else working_part = 0;

if work_status_i eq 1 then working_full = 1;
else working_full = 0;

if work_status_i eq 0 then working_none = 1;
else working_none = 0;

label working_part = "Indicator for working part-time";
label working_full = "Indicator for working full-time";
label working_none = "Indicator for not currently working";

* Creates indicator variables for multiple healthcare coverage;
if insure_new_i eq 2 then insure_two = 1;
else insure_two = 0;

if insure_new_i eq 1 then insure_one = 1;
else insure_one = 0;

if insure_new_i eq 0 then insure_none = 1;
else insure_none = 0;

label insure_two = "Indicator for having 2+ healthcare
coverages";
label insure_one = "Indicator for having 1 healthcare
coverage";
label insure_none = "Indicator for having 0 healthcare
coverage";

* Creates a new variable "age_18_39" that collapses two groups;
* Creates a new variable "age_40_59" that collapses the next two
groups;
age_18_39 = age_18_29 + age_30_39;
age_40_59 = age_40_49 + age_50_59;

* Creates a new variable "working_fullandpart" that collapses two
working groups;
working_fullandpart = working_full + working_part;

```

```
* Creates a new variable "race_nonwhite" that collapses black and
other races;
    race_nonwhite = race_black + race_other;

    label race_nonwhite = "Indicator for non-white race";
run;

*** To check if the binary sex variables were created correctly;
*** Numbers double-checked against Table 1;
*** Results: Numbers match;
proc freq data = puflib.Table3;
    tables sex_female sex_male;
run;

*** To check if the binary race variables were created correctly;
*** Numbers double-checked against Table 1;
*** Results: Numbers match;
proc freq data = puflib.Table3;
    tables race_white race_black race_other;
run;

*** To check if the binary age variables were created correctly;
*** Numbers double-checked against Table 1;
*** Results: Numbers match;
proc freq data = puflib.Table3;
    tables age_18_29 age_30_39 age_40_49 age_50_59 age_60_69;
run;

*** To check if the binary working status variables were created
correctly;
*** Numbers double-checked against Table 1;
*** Results: Numbers match;
proc freq data = puflib.Table3;
    tables working_part working_full working_none;
run;

*** To check if the binary healthcare coverage variables were
created correctly;
*** Numbers double-checked against Table 1;
*** Results: Numbers match;
proc freq data = puflib.Table3;
    tables insure_two insure_one insure_none;
run;

*** Regressions ran after 10/7 meeting with Dr. Lin;
*** Experimental Table 3 - 1st version;
```

```
*** To determine bivariate associations between total costs and
the binary provider type variables;
*** Referent: type_self;
*** Results: No significant association, p = 0.50;
*** Note: Tried making referent type_provider and got negative
estimate;
proc reg data = puflib.Table3;
    model totalcost = type_provider;
run;
quit;

*** To determine bivariate associations between total costs and
the binary sex variables;
*** Referent: sex_male;
*** Results: No significant association, p = 0.96;
*** Note: Tried making referent sex_female and got negative
estimate;
proc reg data = puflib.Table3;
    model totalcost = sex_female;
run;
quit;

*** To determine bivariate associations between total costs and
the binary race variables;
*** Referent: race_nonwhite;
*** Results: No significant association;
proc reg data = puflib.Table3;
    model totalcost = race_white;
run;
quit;

*** To determine bivariate associations between total costs and
the binary age variables;
*** Referent: age_18_29;
*** Results: No significant associations;
*** Note: Parameter estimates are okay here, but earlier code
shows that it will require collapsed age var for the other
outcomes;
proc reg data = puflib.Table3;
    model totalcost = age_30_39 age_40_49 age_50_59 age_60_69;
run;
quit;

proc reg data = puflib.Table3;
    model totalcost = age;
    unit = 5;
run;
```

```
*** To determine bivariate associations between total costs and
the binary age variables;
*** Referent: age_40_59;
*** Results: No significant associations;
proc reg data = puflib.Table3;
    model totalcost = age_18_39 age_60_69;
run;
quit;

proc reg data = puflib.Table3;
    model totalcost = age;
run;
quit;

*** To determine bivariate associations between total costs and
the binary marital variables;
*** Referent: marital_unmarried;
*** Results: No significant association;
*** Note: Tried making referent marital_married and got negative
estimate;
proc reg data = puflib.Table3;
    model totalcost = marital_married;
run;
quit;

*** To determine bivariate associations between total costs and
the binary education variables;
*** Referent: education_nocollege;
*** Results: No significant association;
*** Note: Tried making referent education_college and got
negative estimate;
proc reg data = puflib.Table3;
    model totalcost = education_college;
run;
quit;

*** To determine bivariate associations between total costs and
the binary working status variables;
*** Referent: working_full;
*** Results: No significant association;
proc reg data = puflib.Table3;
    model totalcost = working_none working_part;
run;
quit;
```

```

*** To determine bivariate associations between total costs and
the collapsed working status variables;
*** Referent: working_partandfull;
*** Results: No significant association;
proc reg data = puflib.Table3;
    model totalcost = working_none;
run;
quit;

*** To determine bivariate associations between total costs and
the binary healthcare coverage variables;
*** Referent: insure_none;
*** Results: No significant association;
*** Note: Tried making insure_two the referent and got negative
estimate;
*** Note: Tried making insure_one the referent and got negative
estimate;
proc reg data = puflib.Table3;
    model totalcost = insure_one insure_two;
run;
quit;

proc sort data = puflib.Table1_with_hu_6;
    by abtid;
run;

proc sort data = puflib.Table3_with_hu_6;
    by abtid;
run;

*** Created this dataset on 10.26.11 after learning I have to add
in the unmet needs (hu_6) variables;
data puflib.Table3_with_hu_6;
    merge puflib.Table1_with_hu_6
          puflib.Table2_hospitalization_costs;
    by abtid;

    * Creates indicator variables for provider type;
    if support_group_i eq 0 then type_provider = 1;
    else type_provider = 0;

    if support_group_i eq 1 then type_self = 1;
    else type_self = 0;

    label type_provider = "Indicator for provider referral";
    label type_self     = "Indicator for self referral";

```

```
* Creates indicator variables for sex;
  if sex eq 2 then sex_female = 1;
  else sex_female = 0;

  if sex eq 1 then sex_male = 1;
  else sex_male = 0;

  label sex_female = "Indicator for female gender";
  label sex_male = "Indicator for male gender";

* Creates indicator variables for race;
  if race eq 'White' then race_white = 1;
  else race_white = 0;

  if race eq 'Black' then race_black = 1;
  else race_black = 0;

  if race eq 'Other' then race_other = 1;
  else race_other = 0;

  label race_white = "Indicator for white race";
  label race_black = "Indicator for black race";
  label race_other = "Indicator for other race";

* Creates indicator variables for age;
  if age ge 18 and age le 29 then age_18_29 = 1;
  else age_18_29 = 0;

  if age ge 30 and age le 39 then age_30_39 = 1;
  else age_30_39 = 0;

  if age ge 40 and age le 49 then age_40_49 = 1;
  else age_40_49 = 0;

  if age ge 50 and age le 59 then age_50_59 = 1;
  else age_50_59 = 0;

  if age ge 60 and age le 69 then age_60_69 = 1;
  else age_60_69 = 0;

  label age_18_29 = "Indicator for age 18-29";
  label age_30_39 = "Indicator for age 30-39";
  label age_40_49 = "Indicator for age 40-49";
  label age_50_59 = "Indicator for age 50-59";
  label age_60_69 = "Indicator for age 60-69";

* Creates indicator variables for marital status;
  if marital_status eq 1 then marital_married = 1;
  else marital_married = 0;
```



```

if marital_status eq 0 then marital_unmarried = 1;
else marital_unmarried = 0;

label marital_married    = "Indicator for married status";
label marital_unmarried = "Indicator for unmarried status";

```

```

* Creates indicator variables for education;
  if education eq 1 then education_college = 1;
  else education_college = 0;

  if education eq 0 then education_nocollege = 1;
  else education_nocollege = 0;

  label education_college    = "Indicator for college
education";
  label education_nocollege = "Indicator for less than
college education";

* Creates indicator variables for working status;
  if work_status_i eq 2 then working_part = 1;
  else working_part = 0;

  if work_status_i eq 1 then working_full = 1;
  else working_full = 0;

  if work_status_i eq 0 then working_none = 1;
  else working_none = 0;

  label working_part = "Indicator for working part-time";
  label working_full = "Indicator for working full-time";
  label working_none = "Indicator for not currently working";

* Creates indicator variables for multiple healthcare coverage;
  if insure_new_i eq 2 then insure_two = 1;
  else insure_two = 0;

  if insure_new_i eq 1 then insure_one = 1;
  else insure_one = 0;

  if insure_new_i eq 0 then insure_none = 1;
  else insure_none = 0;

  label insure_two    = "Indicator for having 2+ healthcare
coverages";
  label insure_one    = "Indicator for having 1 healthcare
coverage";
  label insure_none   = "Indicator for having 0 healthcare
coverage";

```

```

* Creates a new variable "age_18_39" that collapses two groups;
* Creates a new variable "age_40_59" that collapses the next two
groups;
    age_18_39 = age_18_29 + age_30_39;
    age_40_59 = age_40_49 + age_50_59;

* Creates a new variable "working_fullandpart" that collapses two
working groups;
    working_fullandpart = working_full + working_part;

* Creates a new variable "race_nonwhite" that collapses black and
other races;
    race_nonwhite = race_black + race_other;

    label race_nonwhite = "Indicator for non-white race";
* Creates a new variable "insure" that collapses insure_one and
insure_two;
    insure = insure_one + insure_two;

    label insure = "Indicator for healthcare coverage";
run;

*****
* BIVARIATE ANALYSES - ALL OUTCOMES VS. PROVIDER TYPE          *
*                                                                 *
*****

*** To determine bivariate associations between total inpatient
hospitalization costs and the binary sex variables;
*** Referent: type_self;
*** Results: No significant association, p = 0.36;
proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = type_provider;
run;
quit;

*** To determine bivariate associations between total provider
encounters costs and the binary provider type variables;
*** Referent: type_self;
*** Results: No significant associations, p = 0.38;
proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = type_provider;
quit;

```

```
*** To determine bivariate associations between other health  
costs and the binary provider type variables;
```

```
*** Referent: type_self;
```

```
*** Results: No significant association, p = 0.51;
```

```
proc reg data = puflib.Table3_with_hu_6;  
    model othhccoststot = type_provider;
```

```
run;
```

```
quit;
```

```
*****  
* BIVARIATE ANALYSES - ALL OUTCOMES VS. SEX *  
* * * * *  
*****
```

```
*** To determine bivariate associations between total inpatient  
hospitalization costs and the binary sex variables;
```

```
*** Referent: sex_male;
```

```
*** Results: No significant association, p = 0.63;
```

```
proc reg data = puflib.Table3_with_hu_6;  
    model hospcosttot = sex_female;
```

```
run;
```

```
quit;
```

```
*** To determine bivariate associations between total provider  
encounters costs and the binary sex variables;
```

```
*** Referent: sex_male;
```

```
*** Results: No significant association, p = 0.96;
```

```
proc reg data = puflib.Table3_with_hu_6;  
    model provallcosttot = sex_female;
```

```
run;
```

```
quit;
```

```
*** To determine bivariate associations between other health  
costs and the binary sex variables;
```

```
*** Referent: sex_male;
```

```
*** Results: Significant association, p = 0.04;
```

```
proc reg data = puflib.Table3_with_hu_6;  
    model othhccoststot = sex_female;
```

```
run;
```

```
quit;
```

```
*****  
* BIVARIATE ANALYSES - ALL OUTCOMES VS. RACE *  
* * * * *  
*****
```

```

*** To determine bivariate associations between total inpatient
hospitalization costs and the binary race variables;
*** Referent: race_nonwhite;
*** Results: No significant association, p = 0.63;
proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = race_white;
run;
quit;

```

```

*** To determine bivariate associations between total provider
encounter costs and the binary race variables;
*** Referent: race_nonwhite;
*** Results: No significant association, p = 0.43;
proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = race_white;
run;
quit;

```

```

*** To determine bivariate associations between other health
costs and the binary race variables;
*** Referent: race_nonwhite;
*** Results: No significant association, p = 0.60;
proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = race_white;
run;
quit;

```

```

*****
* BIVARIATE ANALYSES - ALL OUTCOMES VS. AGE                                     *
*                                                                                       *
*****

```

```

*** To determine bivariate associations between total inpatient
hospitalization costs and the binary age variables;
*** Referent: age_40_59;
*** Results: No significant associations;
proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = age_18_39 age_60_69;
run;
quit;

```

```

*** To determine bivariate associations between total provider
encounters costs and the binary age variables;
*** Referent: age_40_59;
*** Results: No significant associations;
proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = age_18_39 age_60_69;
run;
quit;

```



```

*** To determine bivariate associations between total inpatient
hospitalization costs and the binary education variables;
*** Referent: education_nocollege;
*** Results: No significant association, p = 0.30;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = education_college;
run;
quit;

```

```

*** To determine bivariate associations between total provider
encounters costs and the binary education variables;
*** Referent: education_nocollege;
*** Results: Significant association, p = 0.03;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = education_college;
run;
quit;

```

```

*** To determine bivariate associations between other health
costs and the binary education variables;
*** Referent: education_nocollege;
*** Results: No significant association, p = 0.07;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = education_college;
run;
quit;

```

```

*****
* BIVARIATE ANALYSES - ALL OUTCOMES VS. WORKING STATUS          *
*                                                                 *
*****

```

```

*** To determine bivariate associations between total inpatient
hospitalization costs and the binary working status variables;
*** Referent: working_fullandpart;
*** Results: No significant association, p = 0.25;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = working_none;
run;
quit;

```

```

*** To determine bivariate associations between total provider
encounters costs and the binary working status variables;
*** Referent: working_fullandpart;
*** Results: No significant association, p = 0.46;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = working_none;
run;
quit;

```

```

*** To determine bivariate associations between other health
costs and the binary working status variables;
*** Referent: working_fullandpart;
*** Results: No significant association, p = 0.21;
proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = working_none;
run;
quit;

```

```

*****
* BIVARIATE ANALYSES - ALL OUTCOMES VS. HEALTHCARE COVERAGE      *
*                                                                 *
*****

```

```

*** To determine bivariate associations between total inpatient
hospitalization costs and the binary healthcare coverage
variables;
*** Referent: insure_none;
*** Results: No significant associations;
proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = insure_two insure_one;
run;
quit;

```

```

*** 2.29.12 - Try collapsing for robustness;
*** Look above in puflib.Table3 for the collapsed variable
"insure";
*** Results: No significant associations;
*** Note: Parameter estimate for insure(any insurance) is now
$538.91, and SE is $1,633.87;
proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = insure;
run;
quit;

```

```

*** 2.29.12 - Collapsing;
*** Results: No significant associations;
*** Note: Parameter estimate for insure is now $9,080.12 and SE
is $7,624.02;
proc reg data = puflib.Table3_with_hu_6;
    model totalcost = insure;
run;
quit;

```

```

*** 2.29.12 - Collapsing;
*** Results: No significant associations;
*** Note: Parameter estimate for insure is now $8,541.21 and SE
is $7,627.92;
proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = insure;
run;
quit;
*** 2.29.12 - Collapsing;
*** Results: No significant associations;
*** Note: Parameter estimate for insure is now $77.61 and SE is
$217.21;
proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = insure;
run;
quit;

*** To determine bivariate associations between total provider
encounters costs and the binary healthcare coverage variables;
*** Referent: insure_none;
*** Results: No significant association for insure_one, p = 0.03
for insure_two;
proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = insure_two insure_one;
run;
quit;

*** To determine bivariate associations between other health
costs and the binary healthcare coverage variables;
*** Referent: insure_none;
*** Results: No significant associations;
proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = insure_two insure_one;
run;
quit;

*****
* BIVARIATE ANALYSES - ALL OUTCOMES VS. UNMET NEED          *
*                                                           *
*****

*** To determine bivariate associations between total costs and
the binary unmet need variables;
*** Referent: hu_met_need;
*** Results: p = 0.02, statistically significant association;
proc reg data = puflib.Table3_with_hu_6;
    model totalcost = hu_unmet_need;
run;
quit;

```



```

*** To determine bivariate associations between total inpatient
hospitalization costs and the binary unmet need variables;
*** Referent: hu_met_need;
*** Results: p = 0.10, no significant association;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = hu_unmet_need;
run;
quit;

```

```

*** To determine bivariate associations between total provider
encounter costs and the binary unmet need variables;
*** Referent: hu_met_need;
*** Results: p = 0.06, almost statistically significant
association;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = hu_unmet_need;
run;
quit;

```

```

*** To determine bivariate associations between other health
costs and the binary unmet need variables;
*** Referent: hu_met_need;
*** Results: p = 0.28, no significant association;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = hu_unmet_need;
run;
quit;

```

```

*****
* BIVARIATE ANALYSES - OTC MEDS COLUMN *
* * * * *
*****

```

```

*** Edit: 03.01.12 - To add in OTC meds column for Table 3;
*** Results: p = ;

```

```

proc reg data = puflib.Table3_with_hu_6;
    model otccoststot = type_provider;
    model otccoststot = sex_female;
    model otccoststot = race_white;
    model otccoststot = age_18_39;
    model otccoststot = age_60_69;
    model otccoststot = marital_married;
    model otccoststot = education_college;
    model otccoststot = working_none;
    model otccoststot = insure;
    model otccoststot = hu_unmet_need;
run;

```

```

quit;

*****
* TABLE 4 - ADJUSTED ANALYSES
*
*****

*** Final model to fill in the figures for the total cost column;
*** Results: All aren't significant associations except for
hu_unmet_need ( p = 0.03);
proc reg data = puflib.Table3_with_hu_6;
    model totalcost = type_provider sex_female race_white
    age_18_39 age_60_69 marital_married education_college
    working_none insure_two insure_one hu_unmet_need;
run;
quit;

*** 2.29.12 - Collapsed healthcare coverage version;
proc reg data = puflib.Table3_with_hu_6;
    model totalcost = type_provider sex_female race_white
    age_18_39 age_60_69 marital_married education_college
    working_none insure hu_unmet_need;
run;
quit;

*** Final model to fill in the figures for the inpatient hospital
costs column;
*** Results: No significant associations;
proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = type_provider sex_female race_white
    age_18_39 age_60_69 marital_married education_college
    working_none insure_two insure_one hu_unmet_need;
run;
quit;

*** 2.29.12 - Collapsed healthcare coverage version;
proc reg data = puflib.Table3_with_hu_6;
    model hospcosttot = type_provider sex_female race_white
    age_18_39 age_60_69 marital_married education_college
    working_none insure hu_unmet_need;
run;
quit;

*** Final model to fill in the figures for the total provider
encounter costs column;
*** Results: No significant associations;
proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = type_provider sex_female race_white
    age_18_39 age_60_69 marital_married education_college
    working_none insure_two insure_one hu_unmet_need;
run;
quit;

```

```
*** 2.29.12 - Collapsed healthcare coverage version;
proc reg data = puflib.Table3_with_hu_6;
    model provallcosttot = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure hu_unmet_need;
run;
quit;

*** Final model to fill in the figures for the other health costs
column;
*** Results: No significant associations except for sex_female, p
= 0.03;
proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure_two insure_one hu_unmet_need;
run;
quit;

*** 2.29.12 - Collapsed healthcare coverage version;
proc reg data = puflib.Table3_with_hu_6;
    model othhccoststot = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure hu_unmet_need;
run;
quit;

*** 03.05.12 - Re-ran to include the separated OTC meds column
now;
proc reg data = puflib.Table3_with_hu_6;
    model totalcost = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure hu_unmet_need;
    model hospcosttot = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure hu_unmet_need;
    model provallcosttot = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure hu_unmet_need;
    model otccoststot = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure hu_unmet_need;
    model othhccoststot = type_provider sex_female race_white
age_18_39 age_60_69 marital_married education_college
working_none insure hu_unmet_need;
run;
quit;
```