Paper PO-68

A National Study of Health Services Utilization and Cost of Care with SAS®: Analyses from the 2011 Medical Expenditure Panel Survey

Seungyoung Hwang, Johns Hopkins University Bloomberg School of Public Health

ABSTRACT

This paper provides an in-depth analysis, with the example SAS code, to examine health services utilization and cost of care associated with mood disorders among the older population aged 65 or older in the United States. A cross-sectional study design is used to identify two groups of seniors with mood disorders (n = 441) and without mood disorders (n = 3,822) using the 2011 Medical Expenditure Panel Survey (MEPS). Multivariate regression analyses using the SAS survey analysis procedures are conducted to estimate the incremental health services and direct medical costs (inpatient, outpatient, emergency room, prescription drugs, and other) attributable to mood disorders.

INTRODUCTION

The Medical Expenditure Panel Survey (MEPS) data released from the Agency for Healthcare Research and Quality (AHRQ) are important tools for health care services research and policy in an era of health care reform. Studies using MEPS data found incremental healthcare expenditures for individuals with various conditions (e.g., asthma (1), anxiety disorders (2), hypertension (3), back pain (4), comorbid depression (5), allergic rhinitis (6), rheumatoid arthritis (7), thyroid disorders (8)). But there is a lack of information on healthcare costs and sources of care for older adults with mood disorders in the United States.

The objective of this paper was to develop estimates of the diagnosis of mood disorders among seniors in the U.S. and the association of mood disorders with health services utilization and cost of medical care. We hypothesized that mood disorders would be associated with higher health services utilization and costs, even after adjusting for potentially influential personal and clinical characteristics.

METHODS

STUDY SETTING AND DESIGN

For this research, data were pooled from the 2011 Medical Expenditure Panel Survey Household Component (MEPS-HC) public use files which were the latest available MEPS datasets at the time of study. More specifically, data files used in this study were full year consolidated data file and medical conditions file from MEPS-Household Components. The household component collected data from a nationally representative sample of the civilian non-institutionalized population of the U.S.; and was intended to provide national estimates of healthcare utilization, cost, health insurance coverage, and sources of payment.

PARTICIPANTS

From the total survey respondents in 2011 (n=35,313), 33,622 had positive person weights, and were included in the final analysis. Among them, 4,263 were older adults aged 65 or older. The weighted sample of this study was representative of the approximately 45.5 million US civilian non-institutionalized elderly population.

MEASURES

Mood disorders: Participants were identified as having mood disorder(s) if they had *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) codes of 293.83 (Mood disorder in conditions classified elsewhere), 296 (Episodic mood disorders), 300.4 (Dysthymic disorder), or 311

(Depressive disorder, not elsewhere classified) over the one-year study period.

Health services utilization and cost of care: To estimate direct medical costs we used what is actually paid for healthcare services as a measure of cost, rather than billed charges. Charges are based on the predetermined rate per discharge or procedure without consideration of charitable care, erroneous deductibles or coinsurances, denied claims, and uncollected liability (9). As a result, charges may not represent the actual payment of medical care and less accurate than expenditures for cost estimation purposes. Total direct medical costs were calculated by adding all payments on ambulatory care, physician office visits, emergency room visits, hospital inpatient, and prescription drugs. All costs are reported in 2011 dollars.

Covariates under study: Demographic and socioeconomic characteristics were assessed with standard questions regarding age, gender, race/ethnicity, marital status, level of education attainment, geographic region, metropolitan statistical area (MSA), and health insurance coverage. The poverty status was obtained by a comparison between family income and federal poverty line based on family size and composition. It was defined poor (family income is less than 100% federal poverty line), near poor (100% to less than 125%), low income (125% less than 200%), middle income (200% to less than 400%), and high income (greater than or equal to 400%).

Additional potentially influential covariate included in multivariable models was the D'Hoore adaptation of Charlson comorbidity index (CCI) (10). The 17 comorbidities considered in this index were myocardial infarction, congestive heart failure, peripheral vascular disease, dementia, cerebrovascular disease, chronic pulmonary disease, connective tissue disease, ulcer disease, mild liver disease, hemiplegia, moderate or severe renal disease, diabetes, any tumor, leukemia, lymphoma, moderate or severe liver disease, and metastatic solid tumor.

STATISTICAL ANALYSES

For all analyses we employed a level of statistical significance set at $\alpha = 0.05$, recognizing that tests of statistical significance are approximations that serve as aids to interpretation and inference. SAS version 9.3 was used to carry out analyses (SAS Institute Inc., Cary, North Carolina). The analysis proceeded in two phases, considering the mood disorders as the key variable of interest.

First, we compared demographic, socioeconomic, and clinical characteristics of seniors with and without mood disorders in the study sample. Univariate associations between potential explanatory variables and mood disorders were assessed with SAS/STAT® procedures for survey sampling (e.g., PROC SURVEYFREQ procedure for categorical variables and the PROC SURVEYMEANS procedure for continuous variables) (11).

Second, we carried out multivariate regression analyses using PROC SURVEYREG procedure in SAS in order to estimate the incremental medical services, total expenditures, and type of healthcare expenditures (inpatient, outpatient, office-based visits, emergency room, prescription drug, and other) associated with mood disorders, after controlling for demographic, socioeconomic, and clinical characteristics.

Certain racial and/or ethnic populations, low-income families, and minorities are oversampled in the MEPS in order to improve precision of estimates when examining these groups for particular policy interest. Thus, survey sampling strata and primary sampling units (PSUs) were incorporated into our calculation of weighted mean and standard errors. In addition, all individual statistics were computed based on person weight so that the study sample represented the entire U.S. elderly population in 2011. Analyses were not limited to a subgroup of the population (age ≥ 65). Rather, all analyses were stratified by age (65+ or <65) in order to preserve the entire survey design structure for the program by reading in the entire person-level file.

RESULTS

CHARACTERISTICS OF THE STUDY POPULATION

Compared to seniors without mood disorders, those who have been diagnosed with this condition were

more likely to be in the age group of 65 to 69 (41.70% vs. 37.06%; p = 0.0323), female (67.64% vs. 53.85%; p < 0.0001), white (91.39% vs. 85.45%; p = 0.0027), and poor (11.36% vs. 8.87%; p = 0.0114). Alternatively, they were less likely to be married (48.57% vs. 57.27%; p = 0.0176), uninsured (0.59% vs. 1.68%; p = 0.0376). Also, the mean ± SE of CCI score was significantly higher at 1.73 ± 0.11, for sufferers of mood disorders, compared to seniors without mood disorders at 1.30 ± 0.03.

We also compared the two populations based on the 17 comorbidities used in the D'Hoore adaptation of the CCI. The elderly population with mood disorders, relative to those without, were more likely to have congestive heart failure (4.57% vs. 1.59%; p = 0.0002), dementia (4.75% vs. 1.32%; p < 0.0001), cerebrovascular disease (2.11% vs. 0.56%; p = 0.0012), chronic pulmonary disease (23.11% vs. 10.93%; p < 0.0001), and connective tissue disease (9.10% vs. 5.89%; p = 0.0244).

HEALTHCARE SERVICES ASSOCIATED WITH MOOD DISORDERS

A comparison of mean healthcare services use among seniors with and without mood disorders was conducted. After adjusting for age, gender, race/ethnicity, marital status, education, poverty category, geographic region, MSA, health insurance coverage, and the CCI, seniors with mood disorders had more hospital discharges (0.32 vs. 0.18, p = 0.0004), office-based visits (13.70 vs. 8.94, p < 0.0001), emergency room visits (0.46 vs. 0.29, p = 0.0006), and prescriptions (43.23 vs. 25.94, p < 0.0001) than their counterparts. There were no statistically significant differences in the mean number of outpatient visits.

HEALTHCARE EXPENDITURES ASSOCIATED WITH MOOD DISORDERS

Comparisons of average total, inpatient, outpatient, office-based, emergency room, prescription drug, and other medical expenditures for seniors with and without mood disorders were conducted. Seniors with mood disorders had significantly greater average annual total healthcare expenditures ($14,698) than those without mood disorders ($8,741), after controlling for demographic, socioeconomic, and medical comorbidities. The adjusted annual overall incremental medical expenditure associated with mood disorders was estimated at $5,957 per person (SE: $1,292; p < 0.0001).

Inpatient care, estimated at $1,757 (SE: $1,039; p = 0.0924) accounted for the largest proportion of the overall medical expenditures. Both office-based visits and prescription drugs at $1,442 (SE: $445; p = 0.0014 and SE: $285; p < 0.0001) accounted for the second largest proportion of the overall expenditures, followed by other medical expenses at $970 (SE: $308; p = 0.0018). That is, inpatient, prescription medications, and office-based visits together accounted for approximately 78% of the total incremental cost. Cost of inpatient, outpatient, and emergency room visits were not significantly associated with mood disorders.

CONCLUSION

The prevalence of mood disorders among individuals aged 65 or older in 2011 was estimated at 11.38% (5.17 million persons) and their total direct medical costs were estimated at approximately $81.82 billion in 2011 U.S. dollars.

The presence of mood disorders for older adults has a substantial influence on health services utilization and cost of care in the U.S.

Significant savings associated with mood disorders could be realized by cost effective prescription medications which might reduce the need for subsequent inpatient or office-based visits.

REFERENCE

1. Kamble S,Bharmal M: Incremental direct expenditure of treating asthma in the United States. The Journal of asthma : official journal of the Association for the Care of Asthma 2009; 46:73-80
2. Shirneshan E, Bailey J, Relyea G, et al: Incremental direct medical expenditures associated with anxiety disorders for the U.S. adult population: evidence from the Medical Expenditure Panel Survey. Journal of anxiety disorders 2013; 27:720-727

3. Balu S,Thomas J, 3rd: Incremental expenditure of treating hypertension in the United States. American journal of hypertension 2006; 19:810-816; discussion 817
4. Luo X, Pietrobon R, Sun SX, et al: Estimates and patterns of direct health care expenditures among individuals with back pain in the United States. Spine 2004; 29:79-86
5. Egede LE, Zheng D,Simpson K: Comorbid depression is associated with increased health care use and expenditures in individuals with diabetes. Diabetes care 2002; 25:464-470
6. Bhattacharyya N: Incremental healthcare utilization and expenditures for allergic rhinitis in the United States. The Laryngoscope 2011; 121:1830-1833
7. Kawatkar AA, Jacobsen SJ, Levy GD, et al: Direct medical expenditure associated with rheumatoid arthritis in a nationally representative sample from the medical expenditure panel survey. Arthritis care & research 2012; 64:1649-1656
8. Raval AD,Sambamoorthi U: Incremental Healthcare Expenditures Associated with Thyroid Disorders among Individuals with Diabetes. Journal of thyroid research 2012; 2012:418345
9. Finkler SA: The distinction between cost and charges. Ann Intern Med 1982; 96:102-109
10. D'Hoore W, Bouckaert A,Tilquin C: Practical considerations on the use of the Charlson comorbidity index with administrative data bases. Journal of clinical epidemiology 1996; 49:1429-1433
11. SAS Institute Inc.: SAS/STAT® 9.3 User's Guide. Cary, NC. SAS Institute Inc., 2011

ACKNOWLEDGMENTS

I was first stimulated to pursue studies in applied mathematics and statistics by my mentor and friend, Kyung-Taek Jun. I have learned a great deal from him, and his supportive guidance and inspiration have helped me solve the critical questions raised in my research area. My tremendous debts of gratitude for his contribution are immeasurable. As always, most special thanks to Yun Kyoung Ryu for her advice and encouragement in getting me to finish this paper. She has been incredibly generous with her time in reviewing the draft and making many helpful suggestions. All remaining errors, omissions, and weaknesses are my sole responsibility.

RECOMMENDED READING

In Conference Proceedings:

Analyzing U.S. Healthcare Cost and Use with SAS®

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Seungyoung Hwang, MS, MSE

Department of Mental Health,

Bloomberg School of Public Health

Johns Hopkins University

624 North Broadway

Baltimore, MD 21205

Email: shwang25@jhu.edu

ATTACHMENT

```
************************  BEGINNING OF PROGRAM  ************************;

* set library *;
options ps=66 ls=94 center pageno=1 mprint nodate nofmterr nolabel;
libname MyData "C:\Kai Project\Datasets\MEPS 2011";

* copy to work directory *;
proc datasets;
      copy in=MyData out=work;
      select h146
             h147;
quit;


* ============================  Consolidated  ============================ *;
proc sql;
    create table consol1 as
    select DUPERSID,

           PERWT11F,
           VARPSU,
           VARSTR,

           AGE11X,
           SEX,
           RACEX,
           MARRY11X,
           HIDEG,
           POVCAT11,
           REGION11,
           MSA11,
           UNINS11,
           MCDEV11,
           OPAEV11,
           OPBEV11,
           PRVEV11,

           IPDIS11      as ipvisit,
           OPTOTV11     as opvisit,
           OBTOTV11     as officevisit,
           ERTOT11      as ervisit,
           RXTOT11      as rxvisit,

           IPTEXP11     as ipcost,
           OPTEXP11     as opcost,
           OBVEXP11     as officecost,
           ERTEXP11     as ercost,
           RXEXP11      as rxcost,
           TOTEXP11 - sum(IPTEXP11, OPTEXP11, OBVEXP11, ERTEXP11, RXEXP11)
                        as othercost,
           TOTEXP11     as totalcost

        from h147
        where PERWT11F > 0
    order by DUPERSID;
```

```
quit;
* N = 33,622 *;

data consol2 (drop=OPAEV11 OPBEV11);
    set consol1;

    if   AGE11X >= 64 then subpop = 1;
    else subpop = 0;

    if      64 <= AGE11X <= 69 then AGE11X = 1;
    else if 70 <= AGE11X <= 74 then AGE11X = 2;
    else if 75 <= AGE11X <= 79 then AGE11X = 3;
    else                            AGE11X = 4;

    if      RACEX = 1 then RACEX = 1;
    else if RACEX = 2 then RACEX = 2;
    else                   RACEX = 3;

    if      MARRY11X = 1        then MARRY11X = 1;
    else if MARRY11X in (2,3,4) then MARRY11X = 2;
    else                             MARRY11X = 3;

    if      HIDEG = 1      then HIDEG = 1;
    else if HIDEG in (2,3) then HIDEG = 2;
    else if HIDEG = 4      then HIDEG = 3;
    else if HIDEG in (5,6) then HIDEG = 4;
    else if HIDEG = 7      then HIDEG = 5;
    else                        HIDEG = .;

    if REGION11 = -1 then REGION11 = .;

    if MSA11 = -1 then MSA11 = .;

    if   OPAEV11 = OPBEV11 = 2 then PUBLIC = 2;
    else                            PUBLIC = 1;
run;
* N = 33,622 *;


* ============================  Condition  ============================ *;
proc sql;
    create table condition1 as
    select DUPERSID,
           PERWT11F,
           VARPSU,
           VARSTR,

           CCCODEX,
           ICD9CODX

        from h146
        where PERWT11F > 0
    order by DUPERSID;
quit;
* Record = 105,851 *;

data condition2;
```

```
    set condition1;
    if CCCODEX in ("-1", "-9") then delete;  /* pure missing */
run;
* Record = 103,862 *;

data condition3;
    set condition2;

    if      CCCODEX = "657" then mood = 1;
    else                         mood = 0;
    format mood 1.;

    * Charlson comorbidity index*;
    If   ICD9CODX in ("410", "411") then myoc = 1;
    else                                 myoc = 0;
    format myoc 1.;

    if   ICD9CODX in ("398", "402", "428") then cong = 1;
    else                                        cong = 0;
    format cong 1.;

    if   "440" <= ICD9CODX <= "447" then peri = 1;
    else                                 peri = 0;
    format peri 1.;

    if   ICD9CODX in ("290", "291", "294") then deme = 1;
    else                                        deme = 0;
    format deme 1.;

    if   ICD9CODX in ("430", "431", "432", "433", "435") then cere = 1;
    else                                                      cere = 0;
    format cere 1.;

    if   "491" <= ICD9CODX <= "493" then chro = 1;
    else                                 chro = 0;
    format chro 1.;

    if   ICD9CODX in ("710", "714", "725") then conn = 1;
    else                                        conn = 0;
    format conn 1.;

    if   "531" <= ICD9CODX <= "534" then ulce = 1;
    else                                 ulce = 0;
    format ulce 1.;

    if   ICD9CODX in ("571", "573") then mild = 1;
    else                                 mild = 0;
    format mild 1.;

    if   ICD9CODX in ("342", "434", "436", "437") then hemi = 1;
    else                                               hemi = 0;
    format hemi 1.;

    if   ICD9CODX in ("403", "404") or "580" <= ICD9CODX <= "586"
    then rena = 1;
    else rena = 0;
    format rena 1.;
```

```
    if   ICD9CODX in ("250") then diab = 1;
    else                          diab = 0;
    format diab 1.;

    if   "140" <= ICD9CODX <= "195" then tumo = 1;
    else                               tumo = 0;
    format tumo 1.;

    if   "204" <= ICD9CODX <= "208" then leuk = 1;
    else                               leuk = 0;
    format leuk 1.;

    if   ICD9CODX in ("200", "202", "203") then lymp = 1;
    else                                        lymp = 0;
    format lymp 1.;

    if   ICD9CODX in ("070", "570", "572") then live = 1;
    else                                        live = 0;
    format live 1.;

    if      "196" <= ICD9CODX <= "199" then meta = 1;
    else                                    meta = 0;
    format meta 1.;

run;
* Record = 103,862 *;

proc sql;
    create table condition4 as
    select  DUPERSID,

            max(mood) as mood format=1.,

            max(myoc) as myoc format=1.,
            max(cong) as cong format=1.,
            max(peri) as peri format=1.,
            max(deme) as deme format=1.,
            max(cere) as cere format=1.,
            max(chro) as chro format=1.,
            max(conn) as conn format=1.,
            max(ulce) as ulce format=1.,
            max(mild) as mild format=1.,
            max(hemi) as hemi format=1.,
            max(rena) as rena format=1.,
            max(diab) as diab format=1.,
            max(tumo) as tumo format=1.,
            max(leuk) as leuk format=1.,
            max(lymp) as lymp format=1.,
            max(live) as live format=1.,
            max(meta) as meta format=1.

        from condition3
    group by DUPERSID;
quit;
* N = 25,908 *;
```

```
data condition5;
    set condition4;

    CCI = 1*(myoc + cong + peri + deme + cere + chro + conn + ulce + mild) +
          2*(hemi + rena + diab + tumo + leuk + lymp) +
          3*(live) +
          6*(meta);
    format CCI 2.;
run;
* N = 25,908 *;


* ============================  Merging  ============================ *;
proc sql;
    create table analysis1 as
    select  A.DUPERSID,

            A.PERWT11F,
            A.VARPSU,
            A.VARSTR,

            A.subpop,
            B.mood,

            A.AGE11X,
            A.SEX,
            A.RACEX,
            A.MARRY11X,
            A.HIDEG,
            A.POVCAT11,
            A.REGION11,
            A.MSA11,
            A.UNINS11,
            A.MCDEV11,
            A.PUBLIC,
            A.PRVEV11,

            B.myoc,
            B.cong,
            B.peri,
            B.deme,
            B.cere,
            B.chro,
            B.conn,
            B.ulce,
            B.mild,
            B.hemi,
            B.rena,
            B.diab,
            B.tumo,
            B.leuk,
            B.lymp,
            B.live,
            B.meta,
            B.CCI,

            A.ipvisit,
```

```
            A.opvisit,
            A.officevisit,
            A.ervisit,
            A.rxvisit,

            A.ipcost,
            A.opcost,
            A.officecost,
            A.ercost,
            A.rxcost,
            A.othercost,
            A.totalcost

       from consol2 as A left join condition5 as B on A.DUPERSID = B.DUPERSID
    order by A.DUPERSID;
quit;
* N = 33,622 *;

data analysis2;
    set analysis1;

    if mood = . then mood = 0;

    if myoc = . then myoc = 0;
    if cong = . then cong = 0;
    if peri = . then peri = 0;
    if deme = . then deme = 0;
    if cere = . then cere = 0;
    if chro = . then chro = 0;
    if conn = . then conn = 0;
    if ulce = . then ulce = 0;
    if mild = . then mild = 0;
    if hemi = . then hemi = 0;
    if rena = . then rena = 0;
    if diab = . then diab = 0;
    if tumo = . then tumo = 0;
    if leuk = . then leuk = 0;
    if lymp = . then lymp = 0;
    if live = . then live = 0;
    if meta = . then meta = 0;

    if CCI = . then CCI = 0;

run;
* N = 33,622 *;


* ==============  Characteristics of the Study Population  ============== *;
proc surveyfreq data=analysis2;
    weight PERWT11F;
    cluster VARPSU;
    strata VARSTR;
    tables subpop*mood / row CL nofreq nostd;
run;

proc surveyfreq data=analysis2;
    weight PERWT11F;
```

```
    cluster VARPSU;
    strata VARSTR;
    tables subpop*AGE11X*mood / col CL nofreq nostd chisq;
run;
/*
AGE11X
SEX
RACEX
POVCAT11

MARRY11X
UNINS11
*/

proc surveymeans data=analysis2;
    weight PERWT11F;
    cluster VARPSU;
    strata VARSTR;
    class mood;
    domain subpop;
    model CCI = mood / noint solution vadjust=none;
    lsmeans mood / diff CL;
run;

proc surveyfreq data=analysis2;
    weight PERWT11F;
    cluster VARPSU;
    strata VARSTR;
    tables subpop*myoc*mood / col CL nofreq nostd chisq;
run;
/*
myoc
cong
deme
cere
chro
conn
*/


* =  Healthcare Services & Expenditures Associated with Mood Disorders  = *;
data analysis3;
    set analysis2;
    if mood = 0 then mood = 2;
run;

proc surveyreg data=analysis3;
    weight PERWT11F;
    cluster VARPSU;
    strata VARSTR;
    class mood AGE11X SEX RACEX MARRY11X HIDEG POVCAT11 REGION11 MSA11
          UNINS11 MCDEV11 public PRVEV11;
    domain subpop;
    model totalcost = mood AGE11X SEX RACEX MARRY11X HIDEG POVCAT11 REGION11
                      MSA11 UNINS11 MCDEV11 public PRVEV11 CCI
                      / noint solution vadjust=none;
    lsmeans mood / diff CL;
```

```
run;
/*
ipvisit
officevisit
ervisit
rxvisit
opvisit

totalcost
officecost
rxcost
othercost
ipcost
opcost
ercost
*/


* ============================  Conclusion  ============================ *;
proc surveyfreq data=analysis3;
    weight PERWT11F;
    cluster VARPSU;
    strata VARSTR;
    tables subpop*mood / row CL nofreq nostd;
run;
* 11.38% (5.17 million persons) *;

proc surveymeans data=analysis3 mean clm sum;
    weight PERWT11F;
    cluster VARPSU;
    strata VARSTR;
    var totalcost;
    domain subpop*mood;
run;
* $81.82 billion in U.S. dollars *;

*****************************  END OF PROGRAM  *****************************;
```