CALCULATING MEDICARE HEALTH OUTCOMES SURVEY PERFORMANCE MEASUREMENT RESULTS

William H. Rogers, Ph.D.[a,b]
Barbara Gandek, M.S.[a]
Samuel J. Sinclair, M.Ed.[a]

[a]Health Assessment Lab, Waltham, Massachusetts
[b]The Health Institute, Department of Clinical Care Research, New England Medical Center, Boston, Massachusetts

February 2004

The Centers for Medicare & Medicaid Services' Office of Research, Development, and Information (ORDI) strives to make information available to all. Nevertheless, portions of our files including charts, tables, and graphics may be difficult to read using assistive technology.

Persons with disabilities experiencing problems accessing portions of any file should contact ORDI through e-mail at ORDI_508_Compliance@cms.hhs.gov

INTRODUCTION

The purpose of this document is to provide a detailed outline of the steps utilized for the calculation of Medicare Health Outcomes Survey (HOS) Performance Measurement results. The Performance Measurement results are based on risk adjusted mortality rates, and changes in physical and mental functioning and well being, among living beneficiaries over a two-year period. HOS outcomes are analyzed by calculating the difference between actual and expected results over two years for three major outcome variables: death, change in physical health as measured by the SF-36 Physical Component Summary (PCS) score, and change in mental health as measured by the SF-36 Mental Component Summary (MCS) score. For reporting purposes, death and PCS scores are combined to measure change in physical health. Expected results are adjusted for the case-mix of beneficiaries within a plan. Scores are reported as the percentage of respondents within a plan who were better, same, and worse in physical health and in mental health over the two-year period, after adjustment for case-mix. The modeling and parameter estimates utilized in the calculation of all HOS Performance Measurement results were derived from the original Cohort I analysis.

HOS Performance Measurement analysis is limited to beneficiaries who meet the following analytic criteria:

1. Were age sixty-five or older at the time of completing the baseline survey.

2. Had a baseline survey disposition of M10 (mail, complete survey), M11 (mail, partial complete survey), M31 (mail, break-off), T10 (telephone, complete survey), T11 (telephone, partial complete survey), or T31 (telephone, break-off).

3. Had baseline PCS and MCS scores.

4. Were members of a plan at baseline that remained in HOS at follow-up.

The procedure utilized for the calculation of HOS Performance Measurement results can be summarized into the following seven steps:

1. On the beneficiary-level file, derive variables to be used in calculating expected outcomes.
2. Calculate expected outcomes (for death, PCS, and MCS) for each beneficiary.
3. Calculate actual outcomes (for death, PCS, and MCS) for each beneficiary.
4. Calculate national (system-level) averages for actual outcomes (death, PCS, and MCS).
5. Summarize expected and actual outcomes across all beneficiaries within a plan.
6. Compare actual and expected outcomes to derive plan deviation scores (actual minus expected outcomes) and adjusted plan-level results.
7. Calculate overall *F* statistics for plan comparisons and identify outliers (if warranted based on *F test* results).

Please note the content of this document is specific to the calculation of Cohort III Performance Measurement results. Due to changes in the availability and/or coding of selected variables, modifications to specific items may have occurred. For example, survey date was not captured during the administration of the Cohort I Baseline survey; however, survey date has been collected for each subsequent cohort and is utilized in the methodology outlined here. Caution should be exercised if applying this methodology to HOS data sets other than Cohort III.

STEP 1: Derive Variables To Be Used in Calculations of Expected Outcomes

The expected outcome for a beneficiary is based on that beneficiary's characteristics at baseline. For example, the likelihood of a beneficiary dying over a two-year period is related to factors such as the beneficiary's age and any diseases he or she may have. An expected outcome can be calculated for each beneficiary as a function of his/her characteristics at baseline, and the numerical coefficients that quantify the relationship between those characteristics and the outcome of interest.

Research has been conducted to determine the models that best explain case-mix differences in death and PCS or MCS outcomes, using available data from the Medicare Health Outcomes Survey and CMS administrative files. In addition to the SF-36 Health Survey, the Medicare Health Outcomes Survey includes information about beneficiary sociodemographics (e.g., education, income), self-reported chronic conditions, and additional functional status variables. Sociodemographic information for all beneficiaries (age, gender, race, Medicaid status) is available from CMS administrative data. To calculate expected health outcomes, several variables are coded differently than how they are coded in the original HOS files. An explanation of how the variables are recoded follows below.

Derive Sociodemographic Variables from CMS Files

The following variables should be derived from CMS administrative data included in the HOS files.

VARIABLE	DEFINITION
Age	Age (in years) of the beneficiary at time of baseline survey, calculated from the date the baseline survey was completed (Survdate) and CMS date of birth (Bdate). Calculate BDATE as a SAS date using the SAS MDY function (BDATE=MDY (C3DOB)). Age=((SURVDATE-BDATE)/365.25).
Age75	If beneficiary is age 75 or older, calculate AGE75=AGE-75. Otherwise AGE75=0. Introducing this spline variable allows a bent continuous line fit, with a bend at 75.
Age85	If beneficiary is age 85 or older, calculate AGE85=AGE-85. Otherwise AGE85=0. Introducing this spline variable allows a bent continuous line fit, with a bend at 85.
Female	1 if beneficiary is female (C3GENDER=2), 0 if male (C3GENDER=1).
Agef	Interaction term for age and gender. Calculated by multiplying AGE*FEMALE.
Hblack	1 if beneficiary is black according to CMS (C3RACE=2), 0 if not and C3RACE is not missing.
Hasian	1 if beneficiary is Asian according to CMS (C3RACE=4), 0 if not and C3RACE is not missing.
Hhisp	1 if beneficiary is Hispanic according to CMS (C3RACE=5), 0 if not and C3RACE is not missing.
Medicaid	1 if beneficiary has Medicaid (C3MEDICD=1), 0 if not and C3MEDICD is not missing.
Tele1	1 if beneficiary answered by phone (First digit of C3SRVDSP="T"), otherwise equal to 0.

Derive Sociodemographic Variables from Baseline Survey

Additional sociodemographic variables are derived from HOS data.

VARIABLE	DEFINITION
Black	1 if beneficiary is black according to the HOS survey (C3SV_RAC=3), 0 if not and C3SV_RAC is not missing. Missing if C3SV_RAC=missing.
Asian	1 if beneficiary is Asian according to the HOS survey (C3SV_RAC=2), 0 if not and C3SV_RAC is not missing. Missing if C3SV_RAC=missing.
Hispanic	1 if beneficiary is Hispanic according to the HOS survey (C3HISPAN=1), 0 if not and C3HISPAN is not missing. Missing if C3HISPAN=missing.
Pov2	1 if the respondent reports income less than $20,000 (C3HHINC=1 or 2 or 3), 0 if not. However, this variable should be set to missing if the respondent does not answer (C3HHINC is missing) or responds that he or she does not know (C3HHINC=10).
Hs	1 if the respondent has at least a high school education (12 or more years). Hs is 1 if C3EDUC is 3 or 4 or 5 or 6, and is 0 otherwise, if C3EDUC is not missing. Missing if C3EDUC=missing.
Married	1 if the respondent reports being currently married (C3MARITL=1), and is 0 otherwise, if C3MARITL is not missing. Missing if C3MARITL=missing.
Home	1 if the respondent reports owning, or is in the process of buying, his or her own home (C3HMOWN=1), and is 0 otherwise if C3HMOWN is not missing. Missing if C3HMOWN=missing.

Derive Chronic Condition Variables from Baseline Survey

The thirteen conditions on the chronic condition checklist are recoded to be equal to 1 if the beneficiary reports having the condition, as 0 if the beneficiary reports not having the condition, and as missing if the beneficiary does not answer the question. The four items on treatment for specific types of cancer are recoded as 1 if the beneficiary reports being currently under treatment for the cancer; 0 if the beneficiary reports not being under treatment for that cancer or reports that he or she does not have cancer; and missing otherwise. In addition, four variables are created which are averages of the number of conditions that the beneficiary has within a condition group. See below for additional information on the construction of the condition groups.

VARIABLE	DEFINITION
Highbp	1 if the respondent reports high blood pressure (C3HIGHBP=1), 0 if C3HIGHBP=2.
Angnacad	1 if the respondent reports angina (C3ANGCAD=1), 0 if C3ANGCAD=2.
Chf	1 if the respondent reports congestive heart failure (C3CHF=1), 0 if C3CHF=2.
Ami	1 if the respondent reports myocardial infarction (C3AMI=1), 0 if C3AMI=2.
Othrhart	1 if the respondent reports other heart disease (C3OTHHRT=1), 0 if C3OTHHRT=2.
Stroke	1 if the respondent reports stroke (C3STROKE=1), 0 if C3STROKE=2.
Copd_etc	1 if the respondent reports respiratory disease (C3COPD_E=1), 0 if C3COPD_E=2.
Gi_etc	1 if the respondent reports gastrointestinal disorder (C3GI_ETC=1), 0 if C3GI_ETC=2.
Arthhip	1 if the respondent reports arthritis of hip/knee (C3ATHHIP=1), 0 if C3ATHHIP=2.
Arthhand	1 if the respondent reports arthritis of hand/wrist (C3ATHHAN=1), 0 if C3ATHHAN=2.
Sciatica	1 if the respondent reports sciatica (C3SCIATC=1), 0 if C3SCIATC=2.
Diabetes	1 if the respondent reports diabetes (C3DIABET=1), 0 if C3DIABET=2.
Anycanc	1 if the respondent reports any cancer but skin (C3ANYCAN=1), 0 if C3ANYCAN=2.
Colncanc	1 if the respondent reports current treatment for colon cancer (C3COLNCA=1), 0 if C3COLNCA=2 or C3ANYCAN=2 (whenC3COLNCA ≠ 1).
Brstcanc	1 if the respondent reports current treatment for breast cancer (C3BRSTCA=1), 0 if C3BRSTCA=2 or C3ANYCAN=2 (when C3BRSTCA ≠ 1).
Proscanc	1 if the respondent reports current treatment for prostate cancer (C3PROSCA=1), 0 if C3PROSCA=2 or C3ANYCAN=2 (when C3PROSCA ≠ 1).
Lungcanc	1 if the respondent reports current treatment for lung cancer (C3LUNGCA=1), 0 if C3LUNGCA=2 or C3ANYCAN=2 (when C3LUNGCA ≠ 1).
Sci	The average of **chf**, **anycanc**, **colncanc**, and **lungcanc**. Missing data for any of these variables should not be included in the calculation of the average (e.g., if chf is missing, Sci=(anycanc+colncanc+lungcanc)/3). These are conditions that have fairly large logistic regression coefficients for death.
Mci	The average of **copd_etc**, **diabetes**, **stroke**, and **ami**, averaged over the non-missing variables. These are conditions that have moderate positive logistic regression coefficients for death.
Uci	The average of **brstcanc**, **highbp**, **angnacad**, and **othrhart**, averaged over the non-missing variables. These are conditions that have near-zero logistic regression coefficients for death.
Nci	The average of **gi_etc**, **arthhip**, **arthhand**, **sciatica**, and **proscanc**, averaged over the non-missing variables. These are conditions that have negative logistic regression coefficients for death when functional status variables are included in the model. In other words, these are conditions that do not predict death, but do predict poor functioning. For most conditions, patients with poorer functioning are more likely to die, so poor functioning predicts death. But for some conditions, such as arthritis, their functioning is poor for reasons unlikely to cause death. To make up for this, the coefficients for such conditions are negative.

Variables Describing Functional Status

In addition to the PCS and MCS measures, other SF-36 items and additional questions are used in several of the case-mix adjustment models.

VARIABLE	DEFINITION
Pfadl	A variable scored between 0 and 100 using the 10 SF-36 Physical Functioning items and the 6 ADL items, scored analogously to the SF-36 scales (Likert scoring). 100 is the best possible and 0 is the worst possible physical functioning. See next page for more information on the PFADL scoring algorithm.
C3gh	The general health scale from the SF-36. One hundred is the best possible general health score, 0 is the worst.
C3sf	The social functioning scale from the SF-36. One hundred is the best possible social functioning score, 0 is the worst.
C3cmphth	Comparative health status (question #41). This variable is 1 if the respondent reports excellent health relative to his or her peers, 5 if poor health.
C3mcs98	The baseline Mental Component Summary score calculated using standard half-scale missing data imputation. Calculated using norm-based scoring, with a mean of 50 and standard deviation of 10 in the 1998 US General Population.
C3pcs98	The baseline Physical Component Summary score using standard half-scale missing data imputation. Calculated using norm-based scoring, with a mean of 50 and standard deviation of 10 in the 1998 US General Population.
R3mcs98	The follow-up Mental Component Summary score using standard half-scale missing data imputation. Calculated using norm-based scoring, with a mean of 50 and standard deviation of 10 in the 1998 US General Population.
R3pcs98	The follow-up Physical Component Summary score using standard half-scale missing data imputation. Calculated using norm-based scoring, with a mean of 50 and standard deviation of 10 in the 1998 US General Population.

Scoring the PFADL Scale

The PFADL scale combines the 10 items from the SF-36 Physical Functioning scale with the 6 questions about difficulty in Activities of Daily Living (ADLs) included on the Medicare Health Outcomes Survey. By combining all 16 items together, a scale is created which allows lower levels of physical functioning to be measured than by just the SF-36 Physical Functioning scale alone. The following steps should be followed to create the PFADL scale, which uses the items: C3VIGACT, C3MODACT, C3LIFT, C3CLMBSV, C3CLMBON, C3BEND, C3WLKMI, C3WLKBKS, C3WLK1BK, C3BATHDR, C3DIFBTH, C3DIFDRS, C3DIFEAT, C3DIFCHR, C3DIFWLK, C3DIFTOL.

1. Check all 16 items for out-of-range values. A value is out of range if it is less than 1 or if it is greater than 3. All out-of-range values should be set to missing.

2. Count the number of items that the beneficiary answered. This can range from 0 to 16.

3. If a beneficiary is missing any of the 16 items, impute a value for the missing item(s) using the person-specific mean of the answered items. For example, if a beneficiary answered all items except C3DIFTOL, the mean value of the other 15 items should be used for C3DIFTOL.

4. If the beneficiary answered 8 to 16 items (i.e., if the beneficiary answered at least 50% of the items in the scale), calculate a raw score as the sum of all 16 items (after imputing values for missing data). If the beneficiary answered 0 to 7 items (less than 50%), the raw score is missing.

5. If the raw score is not missing, transform the raw score into a score ranging from 0 (lowest possible score) to 100 (highest possible score) using the following algorithm:

    PFADL = ((Raw score-16) / (48-16)) * 100

STEP 2: Calculate Expected Outcomes for Each Beneficiary

Expected values are calculated for five outcomes: death, PCS same or better, PCS better, MCS same or better, and MCS better. The primary outcomes are "alive and PCS same or better" and "MCS same or better." However, expected and actual outcomes for "alive and PCS better" and "MCS better" are necessary in deriving the percentage of beneficiaries who are better, same, or worse; thus expected values for these two measures are also calculated.

Calculate Expected Probability of Death

The models for death are calculated separately from the models for PCS same or better and the models for PCS better. This is done for two reasons. First, the population on which death can be assessed with reasonable case-mix controls consists of all persons who completed the baseline survey. This is a different sample from those who were followed up. Second, the assessment of death occurs in a different statistical context. Death outcomes come from an independent source, and therefore do not share any correlated error with baseline assessments. It is therefore allowable to use the best possible case-mix model from questionnaire data, including but not limited to baseline SF-36 constructs.

Calculation of the expected probability of death is based on the case-mix of the beneficiary at baseline. The most comprehensive case-mix model includes variables for beneficiary sociodemographics, chronic conditions, and functional status. However, not all beneficiaries have data for all of these variables. To address this issue of missing data, a series of cascading logistic regression models was developed (logistic regression was used because the probability of death is a dichotomous variable where 1=dead and 0=not dead). The most comprehensive case-mix model includes the maximum set of disease variables included on the baseline survey, functional health status data, CMS demographics, and survey sociodemographics. The least comprehensive case-mix model contains CMS demographics only, because these variables are available for the entire analytic sample. To get the best available model for all of the data, the most comprehensive model is used for each beneficiary. Thus, if all of the independent variables are present for Model A, this model is used to calculate an expected probability of death for the beneficiary. If not, then Model B is used if all of the independent variables for Model B are not missing, and so on. Only one model is used for each beneficiary, and an expected probability is calculated for every beneficiary.

There are eight possible death models, each of which has a different set of independent variables. Two-year mortality data are available for all beneficiaries who completed the baseline HOS, whether or not the beneficiary's plan remained in HOS. In the Cohort I analysis, two-year death rates did not vary overall between plans that remained and plans that left managed care over the two-year period. Thus, the full sample of beneficiaries who completed the Cohort I baseline survey was used in the modeling, to obtain parameter estimates based on the largest possible number of cases. The models are theory driven, there were only a limited number of independent variables available from the HOS survey, and the sample size was large. Thus, the full sample was used to develop the parameter estimates, rather than developing and testing the models on split samples. The number of cases in each model and a pseudo-R^2 obtained from the Cohort I analysis are provided below.

The eight death models are:

DEATH MODEL	DESCRIPTION	N[*]	Pseudo-R^{2}[*]
A	Sociodemographic variables (including poverty status), baseline health status measures, and individual chronic conditions.	103,176	0.1713
B	Same model as Model A, without the poverty variable (information on income was missing in many surveys).	125,918	0.1707
C	Same as Model A, but uses disease groupings (e.g., Mci) rather than individual disease coefficients.	116,028	0.1728
D	Same as Model C, without the poverty variable.	142,529	0.1722
E	Same as Model C, but does not include any survey demographics, and the CMS race variable is used instead of the survey race variable.	155,009	0.1740
F	Same as Model E, but substitutes baseline PCS and MCS for the other functioning variables (PFADL, C3GH, C3SF, C3CMPHTH).	159,263	0.1575
G	CMS demographic variables and baseline PCS and MCS only.	161,089	0.1312
H	CMS demographics only.	161,622	0.0685

[*] Obtained from the Cohort I analysis.

Parameter estimates from the regression models are used to calculate the logit probability of death for each beneficiary. This logit probability is then transformed into an expected probability of death, using a standard transformation.

After the best possible model is used to calculate the logit of the probability of death for each beneficiary, the expected probability of death is calculated as follows (where ED=Logit of Probability of Death):

Expected Probability of Death = EXPDEATH = $e^{ED}/(1+e^{ED})$

The variables to be used in each of the eight models, and the parameter estimates needed to calculate the logit probability of death, are presented in Table 1. Where appropriate, the significance level of each of the model specific parameter estimates is indicated.

Table 1
Variables and Parameter Estimates for Eight Death Models

	Model A	Model B	Model C	Model D	Model E	Model F	Model G	Model H
Constant	-5.575606*	-5.479995*	-5.087591*	-5.078024*	-5.198753*	-2.978117*	-2.79009*	-8.103182*
age	0.059085*	0.0587169*	0.0543958*	0.0542354*	0.0537461*	0.0510648*	0.0517747*	0.0742366*
female	-0.8752322#	-0.62861+	-0.7792107#	-0.4675309	-0.7027542#	-0.9061578*	-1.177882*	-1.208054*
age75	0.0061941	0.011573	0.0095808	0.0152497	0.0178543+	0.020995#	0.0163566+	0.021457#
age85	0.0409136*	0.0335235#	0.0379172#	0.0329336#	0.0263124#	0.0282295#	0.0272456*	0.0176281+
agef	0.0035911	-0.0000769	0.0029122	-0.0015829	0.0023317	0.0052511	0.0068117+	0.0086523#
medicaid	0.1286891	0.1662695+	0.0997753	0.13158+	0.2554099*	0.3319504*	0.3435888*	0.7668098*
tele1	0.2896335*	0.2778621*	0.3204558*	0.305184*	0.2856352*	0.3623287*	0.4043633*	
home	-0.0892782#	-0.0895851#	-0.0832811#	-0.0856386#				
pov2	0.0193202		0.0001365					
hs	0.0053484	0.0024538	0.0180647	0.0222926				
married	-0.1054084#	-0.1563856*	-0.0925332#	-0.1416102*				
asian	-0.2518545+	-0.3340543#	-0.2974056+	-0.3765179*				
black	-0.1138535	-0.1145334+	-0.1007439	-0.1105249+				
hispanic	-0.1987376+	-0.2697588*	-0.1958991#	-0.2495292*				
pfadl	-0.0163697*	-0.0164215*	-0.0162912*	-0.0159027*	-0.0159076*			
c3gh	-0.0066906*	-0.0070812*	-0.0083231*	-0.0085637*	-0.0087296*			
c3sf	-0.0054495*	-0.0050047*	-0.0058867*	-0.0055334*	-0.0054748*			
c3cmphth	0.2039281*	0.1945027*	0.2047787*	0.1996191*	0.1978326*			
highbp	-0.0415805	-0.0434913						
angnacad	-0.066195	-0.0716785						
chf	0.5274024*	0.5655592*						
ami	0.1161159+	0.1064888+						
othrhart	0.0296731	0.035137						
stroke	0.1479251*	0.1246192#						
copd_etc	0.2535582*	0.2515284*						
gi_etc	-0.1270888+	-0.1184689+						
arthhip	-0.2923423*	-0.3094484*						
arthhand	-0.1795932*	-0.1636888*						
sciatica	-0.2911476*	-0.2925342*						
diabetes	0.232813*	0.2296315*						
anycanc	0.4968147*	0.4735358*						
colncanc	0.318277#	0.2559645+						
brstcanc	-0.0549007	-0.0468566						
proscanc	-0.344142*	-0.33347*						
lungcanc	1.551023*	1.563668*						
sci			1.945164*	1.949997*	1.896067*	2.004104*		
mci			0.673528*	0.7010161*	0.7254682*	0.9865929*		
nci			-1.270036*	-1.266457*	-1.261842*	-1.326124*		
uci			-0.2003313#	-0.1866082#	-0.1746758#	-0.153813#		
hasian					-0.4570633*	-0.3898549#	-0.3741141#	-0.5430358*
hblack					-0.0930509+	-0.0592814	-0.1120944#	0.0439293
hhisp					-0.3896165*	-0.3308611#	-0.4104027*	-0.2850168#
c3pcs98						-0.05125*	-0.0524192*	
c3mcs98						-0.0283542*	-0.0302379*	

* *p*<0.001   # *p*<0.01   + *p*<0.05

Example of Calculation of Logit Probability of Death Using Model A:

For example, if a beneficiary had data for all of the variables included in Model A at baseline, the calculation of the logit probability of death for that beneficiary would be as follows:

If age ≠ missing and female ≠ missing and age75 ≠ missing and age85 ≠ missing and agef ≠ missing and medicaid ≠ missing and tele1 ≠ missing and home ≠ missing and pov2 ≠ missing and hs ≠ missing and married ≠ missing and asian ≠ missing and black ≠ missing and hispanic ≠ missing and pfadl ≠ missing and c3gh ≠ missing and c3sf ≠ missing and c3cmphth ≠ missing and highbp ≠ missing and angnacad ≠ missing and chf ≠ missing and ami ≠ missing and othrhart ≠ missing and stroke ≠ missing and copd_etc ≠ missing and gi_etc ≠ missing and arthhip ≠ missing and arthhand ≠ missing and sciatica ≠ missing and diabetes ≠ missing and anycanc ≠ missing and colncanc ≠ missing and brstcanc ≠ missing and proscanc ≠ missing and lungcanc ≠ missing then

Logit of Probability of Death=(age*.059085)+(female*-.8752322)+(age75*.0061941)+(age85*.0409136)
    +(agef*.0035911)+(medicaid*.1286891)+(tele1*.2896335)+(home*-.0892782)+
    (pov2*.0193202)+(hs*.0053484)+(married*-.1054084)+(asian*-.2518545)+
    (black*-.1138535)+(hispanic*-.1987376)+(pfadl*-.0163697)+(c3gh*-.0066906)+
    (c3sf*-.0054495)+(c3cmphth*.2039281)+(highbp*-.0415805)+(angnacad*-.066195)+
    (chf*.5274024)+(ami*.1161159)+(othrhart*.0296731)+(stroke*.1479251)+
    (copd_etc*.2535582)+(gi_etc*-.1270888)+(arthhip*-.2923423)+(arthhand*-.1795932)+
    (sciatica*-.2911476)+(diabetes*.232813)+(anycanc*.4968147)+(colncanc*.318277)+
    (brstcanc*-.0549007)+(proscanc*-.344142)+(lungcanc*1.551023)+(-5.575606)

Calculate Expected Probability of PCS Same or Better

The expected probability that a beneficiary's PCS score was the same or better over the two-year follow-up period is calculated for all beneficiaries who have both baseline and follow-up PCS scores. The calculation of the expected probability of PCS same or better is based on the case-mix of the beneficiary at baseline. As with the death models, a series of cascading logistic regression models was developed (logistic regression was used because the probability of PCS same or better is a dichotomous variable where 1=PCS same or better and 0=PCS worse). To get the best available model for all of the data, the most comprehensive model is used for each beneficiary. Thus, if all of the independent variables are present for Model A, this model is used to calculate an expected probability of PCS same or better for the beneficiary. If not, then Model B is used if all of the independent variables for Model B are not missing, and so on. Only one model is used for each beneficiary, and an expected probability is calculated for every beneficiary.

There are three possible PCS same or better models, each of which has a different set of independent variables. The number of cases in each model and a pseudo-R^2 obtained from the Cohort I analysis are provided below. The models are:

PCS SAME OR BETTER MODEL	DESCRIPTION	N[*]	Pseudo-R^{2*}
A	Sociodemographic variables (including poverty status).	52,962	0.0034
B	Same as Model A, without the poverty variable.	64,864	0.0037
C	Only CMS variables (age, gender and Medicaid status).	70,732	0.0029

[*] Obtained from the Cohort I analysis.

Parameter estimates from the regression models are used to calculate the logit probability of PCS same or better for each beneficiary. This logit probability is then transformed into an expected probability of PCS same or better, using a standard transformation.

After the best possible model is used to calculate the logit of the expected probability of PCS same or better, calculate expected PCS same or better from the logit probability, limiting the calculation to cases where baseline and follow-up PCS are non-missing (where EPsb=Logit of Expected PCS Same or Better).

> If c3pcs98 $\neq$ missing and r3pcs98 $\neq$ missing
> then Expected PCS Same or Better = EXPPCSSB = $e^{EPsb}/(1+e^{EPsb})$

The variables to be used in each of the three models, and the parameter estimates needed to calculate the logit probability of PCS same or better, are presented in Table 2. Where appropriate, the significance level of each of the model specific parameter estimates is indicated.

Table 2
Variables and Parameter Estimates for Three PCS Same or Better Models

	Model A	Model B	Model C
Constant	3.267278*	3.412188*	3.589634*
age	-0.0338548*	-0.0369439*	-0.0373943*
female	-0.8603157*	-0.7920578*	-0.7698802*
age75	0.0193157#	0.0210253#	0.0193292#
age85	0.0184556	0.0186742	0.0157097
agef	0.0120167*	0.0112148*	0.0107632*
medicaid	0.1061828	0.0257059	-0.0039739
home	0.0397448	0.0359959	
pov2	-0.0832665*		
hs	0.0825153*	0.1141017*	
married	0.0158446	0.0372333	
asian	0.2015847#	0.16932#	
black	0.0174399	0.0140867	
hispanic	-0.0552064	-0.0687168	

$* p<0.001 \quad \# p<0.01 \quad + p<0.05$

The coefficients of age are negative in the PCS same or better model because older patients decline physically faster than younger ones. Recall that same or better is equal to not worse, so the probability of PCS same or better is equal to 1 minus the probability of PCS worse. The relationship between age and PCS same or better has been seen consistently in the Medical Outcomes Study and studies in the Veterans Administration.

Example of Calculation of Logit Probability of PCS Same or Better Using Model A:

For example, if a beneficiary had data for all of the variables included in Model A at baseline, the calculation of the logit probability of PCS same or better for that beneficiary would be as follows:

If age $\neq$ missing and female $\neq$ missing and age75 $\neq$ missing and age85 $\neq$ missing and agef $\neq$ missing and medicaid $\neq$ missing and home $\neq$ missing and pov2 $\neq$ missing and hs $\neq$ missing and married $\neq$ missing and asian $\neq$ missing and black $\neq$ missing and hispanic $\neq$ missing then

Logit of Probability of PCS Same or Better=
    (age*-.0338548)+(female*-.8603157)+(age75*.0193157)+(age85*.0184556)+(agef*.0120167)+
    (medicaid*.1061828)+(home*.0397448)+(pov2*-.0832665)+(hs*.0825153)+
    (married*.0158446)+(asian*.2015847)+(black*.0174399)+(hispanic*-.0552064)+(3.267278)

Calculate Expected Probability of PCS Better

The expected probability that a beneficiary's PCS score was better over the two-year follow-up period is calculated for all beneficiaries who have both baseline and follow-up PCS scores. The calculation of the expected probability of PCS better is based on the case-mix of the beneficiary at baseline. As with the death models, a series of cascading logistic regression models was developed (logistic regression was used because the probability of PCS better is a dichotomous variable where 1=PCS better and 0=PCS same or worse). To get the best available model for all of the data, the most comprehensive model is used for each beneficiary. Thus, if all of the independent variables are present for Model A, this model is used to calculate an expected probability of PCS better for the beneficiary. If not, then Model B is used if all of the independent variables for Model B are not missing, and so on. Only one model is used for each beneficiary, and an expected probability is calculated for every beneficiary.

There are three possible PCS better models, each of which has a different set of independent variables. The number of cases in each model and a pseudo-R^2 obtained from the Cohort I analysis are provided below. The models are:

PCS BETTER MODEL	DESCRIPTION	N[*]	Pseudo-R^{2*}
A	Sociodemographic variables (including poverty status).	52,962	0.0023
B	Same as Model A, without the poverty variable.	64,864	0.0020
C	Only CMS variables (age, gender and Medicaid status).	70,732	0.0007

[*] Obtained from the Cohort I analysis.

Parameter estimates from the regression models are used to calculate the logit probability of PCS better for each beneficiary. The logit probability is then transformed into an expected probability of PCS better, using a standard transformation.

The variables to be used in each of the three models, and the parameter estimates needed to calculate the logit probability of PCS better, are presented in Table 3. Where appropriate, the significance level of each of the model specific parameter estimates is indicated.

	Table 3		
	Variables and Parameter Estimates for Three PCS Better Models		
	Model A	*Model B*	*Model C*
Constant	-2.172204*	-2.148534*	-2.281332*
age	0.004964	0.0060665	0.0070831
female	-0.4577375	-0.5560831+	-0.3755389
age75	0.0021139	-0.0014687	-0.0009189
age85	-0.010275	-0.003094	-0.0090416
agef	0.0069375	0.0080432+	0.00546
medicaid	0.1440785	0.113725	0.1877218#
home	-0.042036	-0.0534817	
pov2	0.103442*		
hs	-0.112051*	-0.1346055*	
married	0.1275917*	0.0835041#	
asian	0.1383128	0.1136843	
black	0.1168021+	0.1529066#	
hispanic	0.0845965	0.114252+	

* *p<0.001   # p<0.01   + p<0.05*

After the best possible model is used to calculate the logit of the expected probability of PCS better, calculate expected PCS better from the logit probability, limiting the calculation to cases where baseline and follow-up PCS are non-missing (where EPb=Logit of Expected PCS Better).

If c3pcs98 $\neq$ missing and r3pcs98 $\neq$ missing
then Expected PCS Better = EXPPCSB = $e^{EPb}/(1+e^{EPb})$

The coefficient of age is positive in Table 3, albeit not significant. It is not clear why this is so, although in previous research into several conditions that are amenable to both good or bad impacts of medical care, patients have been shown to have a higher likelihood of becoming both worse and better over time.

Calculate Expected Probability of MCS Same or Better

The expected probability that a beneficiary's MCS score was the same or better over the two-year follow-up period is calculated for all beneficiaries who have both baseline and follow-up MCS scores. The calculation of the expected probability of MCS same or better is based on the case-mix of the beneficiary at baseline. As with the death models, a series of logistic regression models was developed (logistic regression was used because the probability of MCS same or better is a dichotomous variable where 1=MCS same or better and 0=MCS worse). To get the best available model for all of the data, the most comprehensive model is used for each beneficiary. Thus, if all of the independent variables are present for Model A, this model is used to calculate an expected probability of MCS same or better for the beneficiary. If not, then Model B is used if all of the independent variables for Model B are not missing, and so on. Only one model is used for each beneficiary, and an expected probability is calculated for every beneficiary.

There are three possible MCS same or better models, each of which has a different set of independent variables. The number of cases in each model and a pseudo-R^2 obtained from the Cohort I analysis are provided below. The models are:

MCS SAME OR BETTER MODEL	DESCRIPTION	N[*]	Pseudo-R²[*]
A	Sociodemographic variables (including poverty status).	53,151	0.0146
B	Same as Model A, without the poverty variable.	65,103	0.0127
C	Only CMS variables (age, gender, race and Medicaid status).	71,003	0.0094

[*] Obtained from the Cohort I analysis.

Parameter estimates from the regression models are used to calculate the logit probability of MCS same or better for each beneficiary. The logit probability is then transformed into an expected probability of MCS same or better, using a standard transformation.

The variables to be used in each of the three models, and the parameter estimates needed to calculate the logit probability of MCS same or better, are presented in Table 4. Where appropriate, the significance level of each of the model specific parameter estimates is indicated.

Table 4 Variables and Parameter Estimates for Three MCS Same or Better Models			
	Model A	*Model B*	*Model C*
Constant	4.290668*	4.230617*	4.469469*
age	-0.0396762*	-0.0413088*	-0.0422551*
female	-0.3639759	-0.4101988	-0.5653867+
age75	0.0052427	0.0053527	0.0032464
age85	0.0106093	0.0150838	0.0142201
agef	0.0037154	0.0043972	0.0069618+
medicaid	-0.2678399*	-0.2531204*	-0.3292756*
home	0.1015457*	0.1099513*	
pov2	-0.2237441*		
hs	0.2115211*	0.2636095*	
married	-0.2017257*	-0.1268977*	
asian	-0.1026597	-0.0734853	
black	-0.1326165#	-0.1996325*	
hispanic	-0.078652	-0.1342561#	
hblack			-0.2523381*
hhisp			-0.2387025#
hasian			0.004378

* $p<0.001$ \# $p<0.01$ + $p<0.05$

After the best possible model is used to calculate the logit of the expected probability of MCS same or better, calculate expected MCS same or better from the logit probability, limiting the calculation to cases where baseline and follow-up MCS are non-missing (where EMsb=Logit of Expected MCS Same or Better).

> If c3mcs98 ≠ missing and r3mcs98 ≠ missing
> then Expected MCS Same or Better = EXPMCSSB = $e^{EMsb}/(1+e^{EMsb})$

Calculate Expected Probability of MCS Better

The expected probability that a beneficiary's MCS score was better over the two-year follow-up period is calculated for all beneficiaries who have both baseline and follow-up MCS scores. The calculation of the expected probability of MCS better is based on the case-mix of the beneficiary at baseline. As with the death models, a series of cascading logistic regression models was developed (logistic regression was used because the probability of MCS better is a dichotomous variable where 1=MCS better and 0=MCS same or worse). To get the best available model for all of the data, the most comprehensive model is used for each beneficiary. Thus, if all of the independent variables are present for Model A, that model is used to calculate an expected probability of MCS better for the beneficiary. If not, then Model B is used if all of the independent variables for Model B are not missing, and so on. Only one model is used for each beneficiary, and an expected probability is calculated for every beneficiary.

There are three possible MCS better models, each of which has a different set of independent variables. The number of cases in each model and a pseudo-R^2 obtained from the Cohort I analysis are provided below. The models are:

MCS BETTER MODEL	DESCRIPTION	N[*]	Pseudo-R^{2*}
A	Sociodemographic variables (including poverty status).	53,151	0.0079
B	Same as Model A, without the poverty variable.	65,103	0.0066
C	Only CMS variables (age, gender, race and Medicaid status).	71,003	0.0032

[*] Obtained from the Cohort I analysis.

Parameter estimates from the regression models are used to calculate the logit probability of MCS better for each beneficiary. The logit probability is then transformed into an expected probability of MCS better, using a standard transformation.

The variables to be used in each of the three models, and the parameter estimates needed to calculate the logit probability of MCS better, are presented in Table 5. Where appropriate, the significance level of each of the model specific parameter estimates is indicated.

	Model A	Model B	Model C
Table 5			
Variables and Parameter Estimates for Three MCS Better Models			
	Model A	*Model B*	*Model C*
Constant	-2.385757*	-2.319254*	-2.986842*
age	0.0096731	0.0109386+	0.0156096*
female	0.8203816#	0.7174221#	0.6855663#
age75	0.0062538	0.0086711	0.0082538
age85	-0.0018618	-0.0096264	-0.0025787
agef	-0.0102308+	-0.0084135+	-0.0074519+
medicaid	0.148751	0.2036733#	0.324786*
home	-0.0326809	-0.0314016	
pov2	0.210614*		
hs	-0.2363332*	-0.270556*	
married	-0.0906781#	-0.1512992*	
asian	-0.05392	-0.0029232	
black	0.0251279	0.0629019	
hispanic	0.1929396*	0.1814451*	
hblack			0.1435944#
hhisp			0.2789385*
hasian			0.0787052

* $p<0.001$ # $p<0.01$ + $p<0.05$

After the best possible model is used to calculate the logit of the expected probability of MCS better, calculate expected MCS better from the logit probability, limiting the calculation to cases where baseline and follow-up MCS are non-missing (where EMb=Logit of Expected MCS Better).

If c3mcs98 $\neq$ missing and r3mcs98 $\neq$ missing
then Expected MCS Better = EXPMCSB = $e^{EMb}/(1+e^{EMb})$

STEP 3: Calculate Actual Outcomes for Each Beneficiary

Identify Beneficiaries Who Died Within Two Years of Completing the Baseline Survey

A person is defined as dead at follow-up if he or she died within two years of the baseline survey date and did not fill out a follow-up questionnaire. The date of death for HOS beneficiaries who died was determined from CMS data. The baseline survey date plus two years was used to identify beneficiaries who died within two years of completing the baseline survey. There were a small number of beneficiaries who completed the follow-up survey and then died within the two-year window. These beneficiaries were coded as alive and were included in the PCS/MCS analysis. To determine the beneficiaries who died, first determine if the date of death falls within the two-year window. If the beneficiary died within the two-year window and does not have a follow-up PCS score, code the beneficiary as dead (ACTDEATH=1). If the beneficiary completed a follow-up survey and a PCS score can be calculated for the beneficiary (R3PCS98 is not equal to missing), do not code the beneficiary as dead; ACTDEATH should be set to zero. If the beneficiary died after the two-year window, do not code them as dead; ACTDEATH should be set to zero. Finally, if the beneficiary has not died, code them as living (ACTDEATH=0);

Calculate Actual Change in PCS and MCS Scores

A number of studies have demonstrated that health status scores tend to be higher when the SF-36 is administered by interview (in person or over the telephone) than when the SF-36 is self-administered. This is particularly true for mental health scores (e.g., MCS), in which respondents may be less likely to admit to depressed or anxious feelings in an interview setting. To quantify these differences in HOS, a comparison was done between HOS and Veterans Administration (VA) data. A subset of HOS beneficiaries completed the SF-36 Health Survey as part of an ongoing VA survey that was administered around the same time as the baseline Cohort I Health Outcomes Survey. Comparisons of matched surveys (HOS and VA) for the same person differing by mode of administration (mail or telephone) showed that on average beneficiaries' answers were 1.9 points higher on PCS and 4.5 points higher on MCS when the survey was administered by telephone. As a result, the effect of telephone administration was adjusted by subtracting 1.9 points from all PCS scores and 4.5 points from all MCS scores if the mode of administration was telephone. This was done for both the baseline and follow-up scores.

To do this, first create a variable that indicates if the follow-up survey was administered by telephone:

R3TELE=1 if R3SRVDSP begins with "T", else R3TELE=0

Then adjust the PCS and MCS scores for the impact of telephone administration:

C3PCS1=C3PCS98 – 1.9*TELE1
C3MCS1=C3MCS98 – 4.5*TELE1
R3PCS1=R3PCS98 – 1.9*R3TELE
R3MCS1=R3MCS98 – 4.5*R3TELE

After adjusting for telephone administration, calculate the actual change in PCS and MCS scores using the logic that change equals follow-up score minus baseline score:

PCSDIFF=R3PCS1-C3PCS1
MCSDIFF=R3MCS1-C3MCS1

Determine if Beneficiary was Same or Better

Next, for each beneficiary calculate whether their PCS was better, the same, and better or the same over the two-year period, and whether their MCS was better, the same, and better or the same over the two-year period. PCS is considered to be the same if it changes by less than 5.66 points (plus or minus). MCS is considered to be the same if it changes by less than 6.72 points (plus or minus). PCS is better if it improves by 5.66 points or more, while MCS is better if it improves by 6.72 points or more. The PCS calculation is limited to beneficiaries who have PCS scores both at baseline and follow-up; the MCS calculation is limited to beneficiaries who have MCS scores both at baseline and follow-up.

If the beneficiary had a PCS score at baseline and a PCS score at follow-up, create 0/1 flags that designate if the beneficiary's PCS was better at follow-up, and if the PCS was the same or better at follow-up:

      IF PCSDIFF ≠ missing then:
      IF PCSDIFF Greater than or Equal to 5.66 then ACTPCSB=1; else ACTPCSB=0;
      IF PCSDIFF Greater than or Equal to –5.66 then ACTPCSSB=1; else ACTPCSSB=0.
      IF PCSDIFF=missing, then ACTPCSB=missing and ACTPCSSB=missing.

where ACTPCSB = Actual PCS Better and ACTPCSSB = Actual PCS Same or Better

If the beneficiary had an MCS score at baseline and an MCS score at follow-up, create 0/1 flags that designate if the beneficiary's MCS is better at follow-up, and if the MCS is the same or better at follow-up

      IF MCSDIFF ≠ missing then:
      IF MCSDIFF Greater than or Equal to 6.72 then ACTMCSB=1; else ACTMCSB=0;
      IF MCSDIFF Greater than or Equal to –6.72 then ACTMCSSB=1; else ACTMCSSB=0;
      IF MCSDIFF=missing, then ACTMCSB=missing and ACTMCSSB=missing.

where ACTMCSB = Actual MCS Better and ACTMCSSB = Actual MCS Same or Better.

STEP 4: Calculate National Level Averages for Actual Outcomes

At the national level, calculate the mean actual rates for death, PCS, and MCS across all beneficiaries, using the following variables:

Death: Actual Death (SYSADETH)

PCS: Actual PCS Same or Better (SYSAPSB)
        Actual PCS Better (SYSAPB)

MCS: Actual MCS Same or Better (SYSAMSB)
        Actual MCS Better (SYSAMB)

Use the standard equation for deriving the mean:

$$\text{Mean (X)} = \frac{\Sigma X_i}{N}$$

where ΣX_i is the sum of $X_1 + X_2 + \ldots + X_n$ (where X represents each beneficiary's value for the specified variable), and N is the total number of beneficiaries.

For example, the mean for the actual probability of death is calculated as follows:

$$\text{Mean (SYSADETH)} = \frac{\Sigma \text{SYSADETH}}{N}$$

Missing data should not be included when calculating mean actual values.

STEP 5: Summarize Actual and Expected Results Across Beneficiaries Within a Plan

Within a health plan, calculate the mean expected and actual rates for death, PCS, and MCS across all beneficiaries, using the following variables:

Death: Expected probability of death (EXPDEATH)
       Actual Death (ACTDEATH)

PCS: Actual PCS Same or Better (ACTPCSSB)
       Expected PCS Same or Better (EXPPCSSB)
       Actual PCS Better (ACTPCSB)
       Expected PCS Better (EXPPCSB)

MCS: Actual MCS Same or Better (ACTMCSSB)
       Expected MCS Same or Better (EXPMCSSB)
       Actual MCS Better (ACTMCSB)
       Expected MCS Better (EXPMCSB)

Use the standard equation for deriving the mean:

$$\text{Mean }(X) = \frac{\Sigma\, X_i}{N}$$

where $\Sigma\, X_i$ is the sum of $X_1 + X_2 + \ldots + X_n$ (where X represents each beneficiary's value for the specified variable), and N is the number of beneficiaries within a plan.

For example, the mean for the expected probability of death is calculated as follows:

$$\text{Mean (EXPDEATH)} = \frac{\Sigma\,\text{EXPDEATH}}{N}$$

Missing data should not be included when calculating mean expected and actual values.

STEP 6: Derive Adjusted Plan-Level Results

There are six steps involved in deriving the adjusted plan-level results and associated statistics:

1. Combine plan-level death and PCS results.

2. Calculate plan-level deviation scores.

3. Derive variables needed to calculate adjusted plan-level percentages for better, same, and worse.

4. Calculate adjusted plan-level percentages for better, same, and worse.

5. Calculate the standard error for plan-level outcome variables.

6. Calculate *t* statistics for plan-level outcomes.

1. ***Combine plan-level death and PCS results***. The physical outcome measure used in HOS combines death and change in PCS into two measures of change in physical health: alive and PCS same or better, and alive and PCS better. The following formulas are used to combine the plan-level actual and expected death and PCS data that was calculated in Step 5:

    A. Plan-Level Expected Alive and PCS Same or Better =
        (1-Expected Plan-Level Death Rate) * Expected Plan-Level PCS Same or Better

    B. Plan-Level Actual Alive and PCS Same or Better =
        (1-Actual Plan-Level Death Rate) * Actual Plan-Level PCS Same or Better

    C. Plan-Level Expected Alive and PCS Better =
        (1-Expected Plan-Level Death Rate) * Expected Plan-Level PCS Better

    D. Plan-Level Actual Alive and PCS Better =
        (1-Actual Plan-Level Death Rate) * Actual Plan-Level PCS Better

2. ***Calculate plan-level deviation scores.*** Next, calculate plan-level deviation scores (actual minus expected) for PCS alive and same or better, PCS alive and better, MCS same or better, and MCS better:

    A. Plan-Level Alive and PCS Same or Better Deviation =
        Actual (Plan-Level Alive and PCS Same or Better) –
        Expected (Plan-Level Alive and PCS Same or Better)

    B. Plan-Level Alive and PCS Better Deviation =
        Actual (Plan-Level Alive and PCS Better) – Expected (Plan-Level Alive and PCS Better)

    C. Plan-Level MCS Same or Better Deviation =
        Actual (Plan-Level MCS Same or Better) – Expected (Plan-Level MCS Same or Better)

    D. Plan-Level MCS Better Deviation =
        Actual (Plan-Level MCS Better) – Expected (Plan-Level MCS Better)

3. ***Derive variables needed to calculate adjusted plan-level percentages better, same and worse.*** For reporting purposes, adjusted plan-level percentages are needed rather than deviation scores. The six percentages to be reported are: PCS alive and better, PCS alive and same, PCS worse (including deaths), MCS better, MCS same, and MCS worse. These six percentages can be calculated with four data points for each plan:

>   Plan-Level Alive and PCS Same or Better
>   Plan-Level Alive and PCS Better
>   Plan-Level MCS Same or Better
>   Plan-Level MCS Better

For example, MCS same is equal to MCS same or better minus MCS better. MCS worse is equal to 1 minus MCS same or better.

The general formula for calculating an adjusted plan percentage is as follows:

>   Logit (Plan Deviation) = Logit (Actual Plan Outcome) – Logit (Expected Plan Outcome)

>   Logit (Adjusted Plan Percentage) = Logit (System-Level Outcome) + Logit (Plan Deviation)

>   Adjusted Plan Percentage = $1/(1+e^{(-\text{Logit(Adjusted Plan Percentage)})})$

where logit(x) = ln(x/(1-x)) and ln=natural logarithm.

The national (system-level) outcomes were calculated across all Cohort III plans. A table which shows the actual national (system-level) values for each of the four outcomes of interest, and the logit transformation needed for calculating plan-level results, is below.

National (System-Level) Values for Physical and Mental Health Outcomes

OUTCOME	ACTUAL PERCENTAGE	LOGIT VALUE
Alive and PCS Same or Better	0.6710781	0.7130652
Alive and PCS Better	0.1495432	-1.7381885
MCS Same or Better	0.7903085	1.3267860
MCS Better	0.1612676	-1.6488267

Example of Calculation of Plan-Level Adjusted Alive and PCS Same or Better:

An example of how to calculate a plan-level result is given below. For example, suppose the plan-level actual alive and PCS same or better = 0.70 and the expected alive and PCS same or better=0.66. The calculation of the plan-level adjusted alive and PCS same or better is:

Logit (Alive and PCS Same or Better Deviation)   = ln(0.70/(1-0.70)) – ln(0.66/(1-0.66))
     = ln(2.333333) – ln(1.941176)
     = 0.847298 – 0.663294
     = 0.184004

Logit (Adjusted Plan Percentage Alive and PCS Same or Better) = 0.7130652+ 0.184004 = 0.8970692

Plan-Level Adjusted Alive and PCS Same or Better = $1/(1+e^{(-0.8970692)})$ = 0.7103469

The formulas for calculating the plan-level adjusted alive and PCS same or better, adjusted alive and PCS better, adjusted MCS same or better, and adjusted MCS better follow:

A. *Adjusted Alive and PCS Same or Better*

Logit (Alive and PCS Same or Better Deviation) =
     Logit (Actual Alive and PCS Same or Better) – Logit (Expected Alive and PCS Same or Better)

Logit (Adjusted Plan Percentage Alive and PCS Same or Better) =
     0.7130652 + Logit(Alive and PCS Same or Better Deviation)

Adjusted Alive and PCS Same or Better =
     $1/(1+e^{(-Logit(Adjusted\ Plan\ Percentage\ Alive\ and\ PCS\ Same\ or\ Better))})$

B. *Adjusted Alive and PCS Better*

Logit (Alive and PCS Better Deviation) =
     Logit (Actual Alive and PCS Better) – Logit (Expected Alive and PCS Better)

Logit (Adjusted Plan Percentage Alive and PCS Better) =
     -1.7381885 + Logit(Alive and PCS Better Deviation)

Adjusted alive and PCS better =
     $1/(1+e^{(-Logit(Adjusted\ Plan\ Percentage\ Alive\ and\ PCS\ Better))})$

C. *Adjusted MCS Same or Better*

Logit (MCS Same or Better Deviation) =
   Logit (Actual MCS Same or Better) – Logit (Expected MCS Same or Better)

Logit (Adjusted Plan Percentage MCS Same or Better) =
   1.3267860 + Logit(MCS Same or Better Deviation)

Adjusted MCS Same or Better =
   $1/(1+e^{(-\text{Logit(Adjusted Plan Percentage MCS Same or Better)})})$

D. *Adjusted MCS Better*

Logit (MCS Better Deviation) =
   Logit (Actual MCS Better) – Logit (Expected MCS Better)

Logit (Adjusted Plan Percentage MCS Better) =
   -1.6488267 + Logit(MCS Better Deviation)

Adjusted MCS Better =
   $1/(1+e^{(-\text{Logit(Adjusted Plan Percentage MCS Better)})})$

4. ***Calculate the adjusted percent better, same, and worse for each plan.*** The plan-level adjusted PCS better, PCS same, PCS worse, MCS better, MCS same, and MCS worse can be calculated as follows:

   Adjusted PCS Better = Adjusted Alive and PCS Better
   Adjusted PCS Same = Adjusted Alive and PCS Same or Better – Adjusted Alive and PCS Better
   Adjusted PCS Worse = 1 – Adjusted Alive and PCS Same or Better

   Adjusted MCS Better = Adjusted MCS Better
   Adjusted MCS Same = Adjusted MCS Same or Better – Adjusted MCS Better
   Adjusted MCS Worse = 1 – Adjusted MCS Same or Better

5. ***Calculate the standard error for plan-level outcome variables.*** In order to calculate the standard error for alive and PCS same or better, and alive and PCS better, an approximation can be calculated based on the data. The simplest approximation for the standard error of alive and PCS same or better, and for alive and PCS better, is based on the actual data only, taking each expected value as if it were a constant. However, to avoid binomial instability, the expected values are substituted for the actual values in the calculation. This gives standard errors of :

   Death: Square root of [(Expected Death Rate * (1-Expected Death Rate)) / N for Death Analysis]
   PCS: Square root of [(PCS Rate * (1-PCS Rate)) / N for PCS Analysis]

   Putting these together using propagation of error gives:

A. *Standard Error Alive and PCS Same or Better* =

   Square root of [(((Expected PCS Same or Better)^2*Expected Death Rate*(1-Expected Death Rate))
      / N for Death Analysis) +
   (((1-Expected Death Rate)^2*Expected PCS Same or Better*(1-Expected PCS Same or Better)) /
      N for PCS Analysis)]

B. *Standard error Alive and PCS Better =*

Square root of: [(((Expected PCS Better)^2*Expected Death Rate*(1-Expected Death Rate)) /
    N for Death Analysis) +
(((1-Expected Death Rate)^2*Expected PCS Better*(1-Expected PCS Better)) / N for PCS
Analysis)]

C. *Standard Error MCS Same or Better =*

Square root of [(Expected MCS Same or Better* (1-Expected MCS Same or Better))/N for MCS
    Analysis]

D. *Standard Error of MCS Better =*

Square root of [(Expected MCS Better*(1-Expected MCS Better))/N for MCS Analysis]

6. ***Calculate t statistics for plan-level outcomes.*** To determine the statistical significance of the plan
    deviations, calculate *t* statistics, which are equal to the plan deviation divided by the plan standard
    error. The major outcomes evaluated in the HOS are:

A. *t for Alive and PCS Same or Better =*
        Alive and PCS Same or Better Deviation / Standard Error Alive and PCS Same or Better

B. *t for MCS Same or Better =* MCS Same or Better Deviation / Standard Error MCS Same or Better

STEP 7: Calculate Overall F Statistics for Plan Comparisons and Identify Outliers (if warranted based on F test results)

Overall F statistics for plan comparisons are calculated to determine if the plans differed significantly at the plan-level across all plans. The following three steps need to be performed.

Calculate Residual Values at the Beneficiary Level

A residual value (or difference score) is calculated for each beneficiary by subtracting the expected value from the actual value for each of the designated outcomes.

> Residual Death = (Actual Death) – (Expected Death)
> Residual PCS Same or Better = (Actual PCS Same or Better) – (Expected PCS Same or Better)
> Residual PCS Better = (Actual PCS Better) – (Expected PCS Better)
> Residual MCS Same or Better = (Actual MCS Same or Better) – (Expected MCS Same or Better)
> Residual MCS Better = (Actual MCS Better) – (Expected MCS Better)

Calculate Overall F Statistics

Overall F statistics are calculated using ANOVA, where plan is the independent variable and each residual value is the dependent variable. Overall F statistics were calculated for each of the five designated outcomes; however, the two primary outcomes specified a priori for the HOS analysis are "alive and PCS same or better" and "MCS same or better." P values less than or equal to 0.05 indicate that the plans differed significantly on the outcome of interest.

Identification of Outlier Plans

If the F statistics reveal that the plans differed significantly, outlier plans are identified. A t statistic, expressing the significance of the plan differences from the average national results, was calculated in Step 6 by dividing the plan deviation by the standard error of the deviation. Plans with a t statistic ≥ 2 are designated as significantly better than expected, while plans with a t statistic ≤ -2 are designated as significantly worse than expected, compared to the average national results.

TECHNICAL ASSISTANCE

The Medicare Health Outcomes Survey Information and Technical Support Telephone Line (1-888-880-0077), as well as the HOS e-mail address (hos@azqio.sdps.org), are available to provide assistance with questions. Additionally, the Medicare HOS website provides responses to Frequently Asked Questions (*http://www.cms.hhs.gov/surveys/hos*).

APPENDIX A

```
/******************************************************************************/
/* C3NCSMIX+FTEST.sas                                                        */
/*                                                                           */
/* Purpose:                                                                  */
/* Produce case-mix adjusted scores for Cohort III Performance Measurement   */
/*                                                                           */
/* Created by HAL, 9/28/01                                                   */
/* F Test Added by HSAG, 5/16/02 based on code provided by HAL               */
/* Adapted by HSAG, 7/12/02, for the C2PMR                                   */
/* Adapted by HSAG, 5/28/03, for the C3PMR                                   */
/******************************************************************************/


/****************************************************/
/* CONTROL PANEL (Macro Variables)              */
/****************************************************/


*--------------------------------------*
| Define Cohorts                       |
*--------------------------------------*;
%Let C = C3;
%Let R = R3;
%Let P = P3;
*---------------------------------------------*
| Path to library containing INPUT dataset    |
| and LIBNAME                                 |
*---------------------------------------------*;
%Let libpath = C:\D_drive\C3PMR\data;
%Let lib = C3PMR;


*---------------------------------------------*
| Input Data Set                              |
*---------------------------------------------*;
%Let in = c3r3strp;


*---------------------------------------------*
| Output Beneficiary Level Data Sets          |
*---------------------------------------------*;
* From Case-Mix;
%Let pers = c3pmpers;
* From F Test (act - exp);
%Let pers2 = c3pmres;


*---------------------------------------------*
| Output Plan Level Data Set                  |
*---------------------------------------------*;
%Let plan = c3pmplan;


/****************************************************/
/****************************************************/
/****************************************************/
LIBNAME &lib "&libpath";

DATA temp1;
      set &lib..&in
      (keep=&C.pcs98
            &C.mcs98
            &R.pcs98
            &R.mcs98
            &C.edob
            &C.rndnum
            &C.vendor
            &C.srvdsp
            &P.cntrnm
            &p.planst
            &C.age
```

```
                &p.hdod
                &C.gender
                &C.race
                &C.sv_rac
                &C.hispan
                &C.medicd
                &C.hhinc
                &C.educ
                &C.maritl
                &C.hmown
                &C.highbp
                &C.angcad
                &C.chf
                &C.ami
                &C.othhrt
                &C.stroke
                &C.copd_e
                &C.gi_etc
                &C.athhip
                &C.sciatc
                &C.diabet
                &C.anycan
                &C.colnca
                &C.brstca
                &C.prosca
                &C.lungca
                &C.vigact
                &C.modact
                &C.lift
                &C.clmbsv
                &C.clmbon
                &C.bend
                &C.wlkmi
                &C.wlkbks
                &C.wlk1bk
                &C.bathdr
                &C.difbth
                &C.difdrs
                &C.difeat
                &C.difchr
                &C.difwlk
                &C.diftol
                &C.athhan
                &C.cmphth
                &C.svdate
                &R.srvdsp
                &p.cntrnm
                &R.planst
                &C.genhth
                &C.sckesy
                &C.ashlth
                &C.hthwse
                &C.hthext
                &C.soclmt
                &C.sclact
                HICNUM
                &C.patid
        rename=(&p.hdod=&C.ssadod ));

                if &C.svdate='99999999' then &C.svdate=' ';
RUN;

DATA temp;
        Set temp1;
                GH1 = &C.genhth;
                GH2 = &C.sckesy;
                GH3 = &C.ashlth;
                GH4 = &C.hthwse;
                GH5 = &C.hthext;
                SF1 = &C.soclmt;
```

```
                  SF2 = &C.sclact;
RUN;


*******************************************************************************
*******************************************************************************
*  Generating GH and SF (0-100) Scale Scores                                 *
*  Used in death analysis                                                    *
*******************************************************************************
*******************************************************************************;
data input (drop = i -- rawgh rsf1 -- rawsf);
       set temp;

/* Calculating General Health Scale Score */
ARRAY GHP(5) GH1-GH5;

DO I= 1 TO 5;
   IF GHP(I) < 1 OR GHP(I) > 5 THEN GHP(I) = .;
END;

IF GH1 = 1 THEN RGH1 =   5;
IF GH1 = 2 THEN RGH1 = 4.4;
IF GH1 = 3 THEN RGH1 = 3.4;
IF GH1 = 4 THEN RGH1 =   2;
IF GH1 = 5 THEN RGH1 =   1;

RGH3 = 6 - GH3;
RGH5 = 6 - GH5;

GHNUM = N(GH1,GH2,GH3,GH4,GH5);
GHMEAN = MEAN(RGH1,GH2,RGH3,GH4,RGH5);

ARRAY RGH(5) RGH1 GH2 RGH3 GH4 RGH5;

DO I= 1 TO 5;
   IF RGH(I) = . THEN RGH(I) = GHMEAN;
END;

IF GHNUM GE 3  THEN RAWGH = SUM(RGH1,GH2,RGH3,GH4,RGH5);
&C.GH = ((RAWGH - 5)/(25-5)) * 100;

/* Calculating Social Functioning Scale Score */
ARRAY SOC(2) SF1-SF2;

DO I = 1 TO 2;
    IF SOC(I) < 1 OR SOC(I) > 5 THEN SOC(I) = .;
END;

RSF1 = 6 - SF1;
SFNUM = N(SF1,SF2);
SFMEAN = MEAN(RSF1,SF2);

ARRAY RSF(2) RSF1 SF2;

DO I = 1 TO 2;
   IF RSF(I) = . THEN RSF(I) = SFMEAN;
END;

IF SFNUM GE 1 THEN RAWSF = SUM(RSF1,SF2);
&C.SF = ((RAWSF - 2)/(10-2)) * 100;

label &C.GH = "&C. General Health Scale Score"
        &C.SF = "&C. Social Functioning Scale Score";
run;

*******************************************************************************
*******************************************************************************
*  Determining Average Survey Date for Mail & Telephone Surveys based on     *
*  Survey Disposition Code                                                   *
*******************************************************************************
```

```
****************************************************************************;
data avgdt;
        set temp (keep = &C.svdate &C.srvdsp &C.vendor);
        if substr(&C.srvdsp,1,1)="M" then do;
                &C.admin = "M";
        end;
        if substr(&C.srvdsp,1,1)="T" then do;
                &C.admin = "T";
        end;
        label &C.admin = "&C Baseline Mode of Admin";
run;

data avgdt_m;
        set avgdt;
        where &C.admin = "M";
        &C.mdate = mdy(substr(&C.svdate,1,2),substr(&C.svdate,3,2),substr(&C.svdate,5,4));
run;

data avgdt_t;
        set avgdt;
        where &C.admin = "T";
        &C.tdate = mdy(substr(&C.svdate,1,2),substr(&C.svdate,3,2),substr(&C.svdate,5,4));
run;

proc means noprint data=avgdt_m;
        class &C.vendor;
        var &c.mdate;
        output out = m_avg
                MEAN = &C.avgmdt;
run;

data m_avg2;
        set m_avg (keep = &C.vendor &C.avgmdt);
        avgmdt = trim(left(month(&C.avgmdt))) || "/"
                        || trim(left(day(&C.avgmdt))) || "/"
                        || trim(left(year(&C.avgmdt)));
        label &C.avgmdt = "&C Average SAS Mail Date"; /* To be used in calculations */
        label avgmdt = "&C Average Mail Date"; /* Standard date for viewing purposes */
run;

proc means noprint data=avgdt_t;
        class &C.vendor;
        var &c.tdate;
        output out = t_avg
                MEAN = &C.avgtdt;
run;

data t_avg2;
        set t_avg (keep = &C.vendor &C.avgtdt);
        avgtdt = trim(left(month(&C.avgtdt))) || "/"
                        || trim(left(day(&C.avgtdt))) || "/"
                        || trim(left(year(&C.avgtdt)));
        label &C.avgtdt = "&C Average SAS Tele Date"; /* To be used in calculations */
        label avgtdt = "&C Average Tele Date"; /* Standard date for viewing purposes */
run;

proc print data=m_avg2 noobs label;
        var &C.vendor &C.avgmdt avgmdt;
        title 'Average Baseline Mail Date';
run;

proc print data=t_avg2 noobs label;
        var &C.vendor &C.avgtdt avgtdt;
        title 'Average Baseline Tele Date';
run;

data _NULL_;
        set m_avg2;
        where &C.vendor = 1;
        call symput('mdate1',trim(left(&C.avgmdt)));
```

```
run;

data _NULL_;
      set t_avg2;
      where &C.vendor = 1;
      call symput('tdate1',trim(left(&C.avgtdt)));
run;

data _NULL_;
      set m_avg2;
      where &C.vendor = 2;
      call symput('mdate2',trim(left(&C.avgmdt)));
run;

data _NULL_;
      set t_avg2;
      where &C.vendor = 2;
      call symput('tdate2',trim(left(&C.avgtdt)));
run;

data _NULL_;
      set m_avg2;
      where &C.vendor = 3;
      call symput('mdate3',trim(left(&C.avgmdt)));
run;

data _NULL_;
      set t_avg2;
      where &C.vendor = 3;
      call symput('tdate3',trim(left(&C.avgtdt)));
run;

data _NULL_;
      set m_avg2;
      where &C.vendor = 4;
      call symput('mdate4',trim(left(&C.avgmdt)));
run;

data _NULL_;
      set t_avg2;
      where &C.vendor = 4;
      call symput('tdate4',trim(left(&C.avgtdt)));
run;

data _NULL_;
      set m_avg2;
      where &C.vendor = 5;
      call symput('mdate5',trim(left(&C.avgmdt)));
run;

data _NULL_;
      set t_avg2;
      where &C.vendor = 5;
      call symput('tdate5',trim(left(&C.avgtdt)));
run;

data _NULL_;
      set m_avg2;
      where &C.vendor = 6;
      call symput('mdate6',trim(left(&C.avgmdt)));
run;

data _NULL_;
      set t_avg2;
      where &C.vendor = 6;
      call symput('tdate6',trim(left(&C.avgtdt)));
run;
```

```
     ****************************************************************************
     ****************************************************************************
     Step 1: Create new variables to be used in calculating expected values    *
     and limit dataset to cases to be used in calculating plan-level results.   *
     ****************************************************************************
     ****************************************************************************;

     data &lib..&pers;
        set input;

     **********************************************************************
      Assign date of baseline survey if date is missing                 *
      Using Average Mail & Average Telephone Dates calculated above      *
     **********************************************************************;
     if &C.svdate ne '' then do;
                  survdate =
     mdy(substr(&C.svdate,1,2),substr(&C.svdate,3,2),substr(&C.svdate,5,4));
           end;
           else do;
     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "M" and &C.vendor = 1 then do;
                        survdate = &mdate1;
                  end;
     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "T" and &C.vendor = 1 then do;
                        survdate = &tdate1;
                  end;

     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "M" and &C.vendor = 2 then do;
                        survdate = &mdate2;
                  end;
     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "T" and &C.vendor = 2 then do;
                        survdate = &tdate2;
                  end;

     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "M" and &C.vendor = 3 then do;
                        survdate = &mdate3;
                  end;
     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "T" and &C.vendor = 3 then do;
                        survdate = &tdate3;
                  end;

     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "M" and &C.vendor = 4 then do;
                        survdate = &mdate4;
                  end;
     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "T" and &C.vendor = 4 then do;
                        survdate = &tdate4;
                  end;

     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "M" and &C.vendor = 5 then do;
                        survdate = &mdate5;
                  end;
     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "T" and &C.vendor = 5 then do;
                        survdate = &tdate5;
                  end;

     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "M" and &C.vendor = 6 then do;
                        survdate = &mdate6;
                  end;
     If &C.svdate = '' and substr(&C.srvdsp,1,1) = "T" and &C.vendor = 6 then do;
                        survdate = &tdate6;
                  end;
     end;

     ****************************************************
     Set age equal to baseline age (exact calculation)  *
     ****************************************************;

     age = &C.age;
           /* &C.age had been created prior to this program
               using the following formula:
```

```
                    &C.age = (survdate - bdate)/365.25; */

************************************************************************
Limit sample to seniors who filled out the survey, had a baseline PCS     *
or MCS, and had a reporting unit at follow-up.  Seniors were defined as   *
those 65 years old or older                                               *
***********************************************************************;

if age ge 65;
if &C.srvdsp = 'M10' or &C.srvdsp = 'M11' or &C.srvdsp = 'M31' or &C.srvdsp = 'T10'
or &C.srvdsp = 'T11' or &C.srvdsp = 'T31';
if &C.pcs98 ne . or &C.mcs98 ne .;
if &p.cntrnm ne '';


************************************************************************
Create a variable for Death for those people who died in                  *
the two year window following baseline survey administration              *
***********************************************************************;


************************************************************************
Create the two year window for death using the baseline survey            *
administration date and the date of death variables                       *
i.e., Baseline Survey date+730 days (since 2001 was NOT a leap year) *
***********************************************************************;

deadcat='aliv';

if &C.ssadod ne '' then do;
      deathdt = mdy(substr(&C.ssadod,5,2),substr(&C.ssadod,7,2),substr(&C.ssadod,1,4));
      twoyrdt = survdate + 730; /* Two year window does NOT include a leap year,
therefore 730 days */
      if deathdt le twoyrdt then do;
            deadcat = 'dead';
      end;
end;


************************************************************************
Create a four level variable where                                        *
1 = Not dead                                                              *
2 = Dead in 2 yr window following baseline survey completion, and   *
    there is no follow-up pcs and follow-up mcs                           *
3 = Dead in 2 yr window following baseline survey completion, and   *
    there is a follow-up pcs or follow-up mcs                             *
4 = Dead, but outside 2 yr window                                        *
***********************************************************************;

if &C.ssadod = '' then dethvar = 1;
else if deadcat = 'dead' and &R.pcs98 = . and &R.mcs98 = . then dethvar = 2;
else if deadcat = 'dead' and (&R.pcs98 ne . or &R.mcs98 ne .) then dethvar = 3;
else if &C.ssadod ne '' and deadcat = 'aliv' then dethvar = 4;

************************************************************************
Create the variable for actual death if dead within the                   *
two year window and no follow up pcs or mcs score                         *
***********************************************************************;

if dethvar = 2 then actdeath = 1;
else actdeath = 0;

drop deathdt twoyrdt;

************************************************************************
Create spline variables for age75 and age85 using the age variable   *
***********************************************************************;

if age ge 75 then age75 = age-75;
else age75 = 0;
```

```
if age ge 85 then age85 = age-85;
else age85 = 0;


*****************************************************************
Create a variable for female, using the CMS gender variable    *
*****************************************************************;

if &C.gender = 2 then female = 1;
else if &C.gender = . then female = .;
else female = 0;


***********************************************************
Create variables for race using the CMS race variable    *
***********************************************************;

if &C.race = 2 then hblack = 1;
else if &C.race = . then hblack = .;
else hblack = 0;

if &C.race = 4 then hasian = 1;
else if &C.race = . then hasian = .;
else hasian = 0;

if &C.race = 5 then hhisp = 1;
else if &C.race = . then hhisp = .;
else hhisp = 0;


**************************************************************
Create variables for race using the survey race variable    *
**************************************************************;

if &C.sv_rac = 3 then black = 1;
else if &C.sv_rac = . then black = .;
else black = 0;

if &C.sv_rac = 2 then asian = 1;
else if &C.sv_rac = . then asian = .;
else asian = 0;

if &C.hispan = 1 then hispanic = 1;
else if &C.hispan = . then hispanic = .;
else hispanic = 0;


******************************************************************************
Create other demographic variables, using both CMS and survey variables     *
******************************************************************************;

agef = age*female;

if &C.medicd = 1 then medicaid = 1;
else if &C.medicd = . then medicaid = .;
else medicaid = 0;

if &C.hhinc = 1 or &C.hhinc = 2 or &C.hhinc = 3 then pov2 = 1;
else if &C.hhinc = 10 or &C.hhinc = . then pov2 = .;
else pov2 = 0;

if &C.educ = 3 or &C.educ = 4 or &C.educ = 5 or &C.educ = 6 then hs = 1;
else if &C.educ = . then hs = .;
else hs = 0;

if &C.maritl = 1 then married = 1;
else if &C.maritl = . then married = .;
else married = 0;

if &C.hmown = 1 then home = 1;
```

```
else if &C.hmown = . then home = .;
else home = 0;

phon = substr(&C.srvdsp,1,1);
if phon = 'T' then tele1 = 1;
else tele1 = 0;
drop phon;


*******************************************************************************
Create chronic condition variables using survey chronic conition checklist  *
*******************************************************************************;

if &C.highbp = 1 then highbp = 1;
else if &C.highbp = . then highbp = .;
else highbp = 0;

if &C.angcad = 1 then angnacad = 1;
else if &C.angcad = . then angnacad = .;
else angnacad = 0;

if &C.chf = 1 then chf = 1;
else if &C.chf = . then chf = .;
else chf = 0;

if &C.ami = 1 then ami = 1;
else if &C.ami = . then ami = .;
else ami = 0;

if &C.othhrt = 1 then othrhart = 1;
else if &C.othhrt = . then othrhart = .;
else othrhart = 0;

if &C.stroke = 1 then stroke = 1;
else if &C.stroke = . then stroke = .;
else stroke = 0;

if &C.copd_e = 1 then copd_etc = 1;
else if &C.copd_e = . then copd_etc = .;
else copd_etc = 0;

if &C.gi_etc = 1 then gi_etc = 1;
else if &C.gi_etc = . then gi_etc = .;
else gi_etc = 0;

if &C.athhip = 1 then arthhip = 1;
else if &C.athhip = . then arthhip = .;
else arthhip = 0;

if &C.athhan = 1 then arthhand = 1;
else if &C.athhan = . then arthhand = .;
else arthhand = 0;

if &C.sciatc = 1 then sciatica = 1;
else if &C.sciatc = . then sciatica = .;
else sciatica = 0;

if &C.diabet = 1 then diabetes = 1;
else if &C.diabet = . then diabetes = .;
else diabetes = 0;

if &C.anycan = 1 then anycanc = 1;
else if &C.anycan = . then anycanc = .;
else anycanc = 0;

if &C.colnca = 1 then colncanc = 1;
else if &C.colnca = 2 or &C.anycan = 2 then colncanc = 0;
else colncanc = .;

if &C.brstca = 1 then brstcanc = 1;
```

```
else if &C.brstca = 2 or &C.anycan = 2 then brstcanc = 0;
else brstcanc = .;

if &C.prosca = 1 then proscanc = 1;
else if &C.prosca = 2 or &C.anycan = 2 then proscanc = 0;
else proscanc = .;

if &C.lungca = 1 then lungcanc = 1;
else if &C.lungca = 2 or &C.anycan = 2 then lungcanc = 0;
else lungcanc = .;


**************************************************************
Create variables for disease groupings according to        *
regression coefficients                                    *
**************************************************************;

**Large Positive Regression Coefficients**;
sci = mean(chf,anycanc,colncanc,lungcanc);

**Moderate Positive Regression Coefficients**;
mci = mean(copd_etc,diabetes,stroke,ami);

**Near Zero Regression Coefficients**;
uci = mean(brstcanc,highbp,angnacad,othrhart);

**Negative Regression Coefficients**;
nci = mean(gi_etc,arthhip,arthhand,sciatica,proscanc);


***************************************************************************************
Score the PFADL scale.  Scale uses all 10 PF and 6 ADL items and is scored      *
from 0-100, where 0=worse and 100=best possible score.  Scale is scored if at   *
least 8 items are completed (half-scale rule) and person-specific value is      *
used for missing item data.                                                     *
***************************************************************************************;

array pfadli(16) &C.vigact &C.modact &C.lift &C.clmbsv &C.clmbon &C.bend &C.wlkmi
&C.wlkbks
                 &C.wlk1bk &C.bathdr &C.difbth &C.difdrs &C.difeat &C.difchr &C.difwlk
&C.diftol;

do i = 1 to 16;
   if pfadli(i) < 1 or pfadli(i) > 3 then pfadli(i) = .;
end;

pfadlnum = n(&C.vigact,&C.modact,&C.lift,&C.clmbsv,&C.clmbon,&C.bend,&C.wlkmi,&C.wlkbks,
&C.wlk1bk,&C.bathdr,&C.difbth,&C.difdrs,&C.difeat,&C.difchr,&C.difwlk,&C.diftol);
pfadmean =
mean(&C.vigact,&C.modact,&C.lift,&C.clmbsv,&C.clmbon,&C.bend,&C.wlkmi,&C.wlkbks,
&C.wlk1bk,&C.bathdr,&C.difbth,&C.difdrs,&C.difeat,&C.difchr,&C.difwlk,&C.diftol);

do i = 1 to 16;
  if pfadli(i)= . then pfadli(i) = pfadmean;
end;

if pfadlnum ge 8 then rawpfadl =
sum(&C.vigact,&C.modact,&C.lift,&C.clmbsv,&C.clmbon,&C.bend,
&C.wlkmi,&C.wlkbks,&C.wlk1bk,&C.bathdr,&C.difbth,&C.difdrs,
&C.difeat,&C.difchr,&C.difwlk,&C.diftol);
pfadl = ((rawpfadl - 16)/(48-16)) * 100;


drop pfadmean pfadlnum rawpfadl;


***************************************************************************************
***************************************************************************************
Step Two: Calculate expected probablity of death and expected change in         *
PCS/MCS for each case.  Calculate logit probability using parameters from       *
```

```
Cohort 1 logistic regression models.  Then calculate expected value.          *
*******************************************************************************
*******************************************************************************;


*******************************************************************************
*******************************************************************************
Step 2A: Calculate expected probability of death. Best possible (lowest        *
lettered) model is used for each case. First calculate ED, the logit           *
probability of death.   Then calculate expected death from the logit           *
probability  of death.                                                         *
*******************************************************************************
*******************************************************************************;



**********************************************
Death Model A                            *
**********************************************;

if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and tele1 ne . and home ne . and pov2 ne . and hs ne . and
married ne . and asian ne . and black ne . and hispanic ne . and &C.gh ne . and
&C.sf ne . and &C.cmphth ne . and highbp ne . and angnacad ne . and chf ne . and
ami ne . and othrhart ne . and stroke ne . and copd_etc ne . and gi_etc ne . and
arthhip ne . and arthhand ne . and sciatica ne . and diabetes ne . and anycanc ne . and
colncanc ne . and brstcanc ne . and proscanc ne . and lungcanc ne . and pfadl ne .
then

ED = (age*.059085)+(female*-.8752322)+(age75*.0061941)+(age85*.0409136)+
        (agef*.0035911)+(medicaid*.1286891)+(tele1*.2896335)+(home*-.0892782)+
        (pov2*.0193202)+(hs*.0053484)+(married*-.1054084)+(asian*-.2518545)+
        (black*-.1138535)+(hispanic*-.1987376)+(pfadl*-.0163697)+(&C.gh*-.0066906)+
        (&C.sf*-.0054495)+(&C.cmphth*.2039281)+(highbp*-.0415805)+(angnacad*-.066195)+
        (chf*.5274024)+(ami*.1161159)+(othrhart*.0296731)+(stroke*.1479251)+
        (copd_etc*.2535582)+(gi_etc*-.1270888)+(arthhip*-.2923423)+(arthhand*-.1795932)+
        (sciatica*-.2911476)+(diabetes*.232813)+(anycanc*.4968147)+(colncanc*.318277)+
        (brstcanc*-.0549007)+(proscanc*-.344142)+(lungcanc*1.551023)+(-5.575606);



*****************************
Death Model B              *
*****************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and tele1 ne . and home ne . and hs ne . and
married ne . and asian ne . and black ne . and hispanic ne . and &C.gh ne . and
&C.sf ne . and &C.cmphth ne . and highbp ne . and angnacad ne . and chf ne . and
ami ne . and othrhart ne . and stroke ne . and copd_etc ne . and gi_etc ne . and
arthhip ne . and arthhand ne . and sciatica ne . and diabetes ne . and anycanc ne . and
colncanc ne . and brstcanc ne . and proscanc ne . and lungcanc ne . and pfadl ne .
then

ED = (age*.0587169)+(female*-.62861)+(age75*.011573)+(age85*.0335235)+
        (agef*-.0000769)+(medicaid*.1662695)+(tele1*.2778621)+(home*-.0895851)+
        (hs*.0024538)+(married*-.1563856)+(asian*-.3340543)+
        (black*-.1145334)+(hispanic*-.2697588)+(pfadl*-.0164215)+(&C.gh*-.0070812)+
        (&C.sf*-.0050047)+(&C.cmphth*.1945027)+(highbp*-.0434913)+(angnacad*-.0716785)+
        (chf*.5655592)+(ami*.1064888)+(othrhart*.035137)+(stroke*.1246192)+
        (copd_etc*.2515284)+(gi_etc*-.1184689)+(arthhip*-.3094484)+(arthhand*-.1636888)+
        (sciatica*-.2925342)+(diabetes*.2296315)+(anycanc*.4735358)+(colncanc*.2559645)+
        (brstcanc*-.0468566)+(proscanc*-.33347)+(lungcanc*1.563668)+(-5.479995);



*****************************
Death Model C              *
*****************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and tele1 ne . and home ne . and pov2 ne . and hs ne . and
married ne . and asian ne . and black ne . and hispanic ne . and &C.gh ne . and
&C.sf ne . and &C.cmphth ne . and sci ne . and mci ne . and nci ne . and
```

```
uci ne . and pfadl ne .
then

ED = (age*.0543958)+(female*-.7792107)+(age75*.0095808)+(age85*.0379172)+
        (agef*.0029122)+(medicaid*.0997753)+(tele1*.3204558)+(home*-.0832811)+
        (pov2*.0001365)+(hs*.0180647)+(married*-.0925332)+(asian*-.2974056)+
        (black*-.1007439)+(hispanic*-.1958991)+(pfadl*-.0162912)+(&C.gh*-.0083231)+
        (&C.sf*-.0058867)+(&C.cmphth*.2047787)+(sci*1.945164)+(mci*.673528)+
        (nci*-1.270036)+(uci*-.2003313)+(-5.087591);


*****************************
Death Model D              *
*****************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and tele1 ne . and home ne . and hs ne . and
married ne . and asian ne . and black ne . and hispanic ne . and &C.gh ne . and
&C.sf ne . and &C.cmphth ne . and sci ne . and mci ne . and nci ne . and
uci ne . and pfadl ne .
then

ED = (age*.0542354)+(female*-.4675309)+(age75*.0152497)+(age85*.0329336)+
        (agef*-.0015829)+(medicaid*.13158)+(tele1*.305184)+(home*-.0856386)+
        (hs*.0222926)+(married*-.1416102)+(asian*-.3765179)+
        (black*-.1105249)+(hispanic*-.2495292)+(pfadl*-.0159027)+(&C.gh*-.0085637)+
        (&C.sf*-.0055334)+(&C.cmphth*.1996191)+(sci*1.949997)+(mci*.7010161)+
        (nci*-1.266457)+(uci*-.1866082)+(-5.078024);


*****************************
Death Model E              *
*****************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and tele1 ne .  and hasian ne . and hblack ne . and hhisp ne . and
&C.gh ne . and &C.sf ne . and &C.cmphth ne . and sci ne . and mci ne . and nci ne . and
uci ne . and pfadl ne .
then

ED = (age*.0537461)+(female*-.7027542)+(age75*.0178543)+(age85*.0263124)+
        (agef*.0023317)+(medicaid*.2554099)+(tele1*.2856352)+
        (hasian*-.4570633)+(hblack*-.0930509)+(hhisp*-.3896165)+(pfadl*-.0159076)+
        (&C.gh*-.0087296)+(&C.sf*-.0054748)+(&C.cmphth*.1978326)+(sci*1.896067)+
        (mci*.7254682)+(nci*-1.261842)+(uci*-.1746758)+(-5.198753);


*****************************
Death Model F              *
*****************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and tele1 ne .  and hasian ne . and hblack ne . and hhisp ne . and
&C.pcs98 ne . and &C.mcs98 ne . and sci ne . and mci ne . and nci ne . and
uci ne .
then

ED = (age*.0510648)+(female*-.9061578)+(age75*.020995)+(age85*.0282295)+
        (agef*.0052511)+(medicaid*.3319504)+(tele1*.3623287)+
        (hasian*-.3898549)+(hblack*-.0592814)+(hhisp*-.3308611)+
        (&C.pcs98*-.05125)+(&C.mcs98*-.0283542)+(sci*2.004104)+
        (mci*.9865929)+(nci*-1.326124)+(uci*-.153813)+(-2.978117);


*****************************
Death Model G              *
*****************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and tele1 ne .  and hasian ne . and hblack ne . and hhisp ne . and
```

```
&C.pcs98 ne . and &C.mcs98 ne .
then

ED = (age*.0517747)+(female*-1.177882)+(age75*.0163566)+(age85*.0272456)+
        (agef*.0068117)+(medicaid*.3435888)+(tele1*.4043633)+
        (hasian*-.3741141)+(hblack*-.1120944)+(hhisp*-.4104027)+
        (&C.pcs98*-.0524192)+(&C.mcs98*-.0302379)+(-2.79009);


*****************************
Death Model H                    *
*****************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne .  and hasian ne . and hblack ne . and hhisp ne .
then

ED = (age*.0742366)+(female*-1.208054)+(age75*.021457)+(age85*.0176281)+
        (agef*.0086523)+(medicaid*.7668098)+(hasian*-.5430358)+(hblack*.0439293)+
        (hhisp*-.2850168)+(-8.103182);


***********************************************************************
Calculate expected death from the logit probablity of death.        *
***********************************************************************;

ExpDeath = exp(ED)/(1+exp(ED));


******************************************************************************************
******************************************************************************************
Step 2B: Calculate expected PCS same or better. Best possible (lowest lettered)        *
model is used for each case. First calculate EPsb, the logit probablity of PCS same    *
or better.  Then calculate expected PCS same or better from the logit.                 *
******************************************************************************************
******************************************************************************************;


**********************************
PCS Same+Better Model A           *
**********************************;

if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and pov2 ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EPsb = (age*-.0338548)+(female*-.8603157)+(age75*.0193157)+(age85*.0184556)+
        (agef*.0120167)+(medicaid*.1061828)+(home*.0397448)+
        (pov2*-.0832665)+(hs*.0825153)+(married*.0158446)+
        (asian*.2015847)+(black*.0174399)+(hispanic*-.0552064)+(3.267278);


**********************************
PCS Same+Better Model B           *
**********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EPsb = (age*-.0369439)+(female*-.7920578)+(age75*.0210253)+(age85*.0186742)+
        (agef*.0112148)+(medicaid*.0257059)+(home*.0359959)+
        (hs*.1141017)+(married*.0372333)+
        (asian*.16932)+(black*.0140867)+(hispanic*-.0687168)+(3.412188);


**********************************
PCS Same+Better Model C           *
```

```
**********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne .
then

EPsb = (age*-.0373943)+(female*-.7698802)+(age75*.0193292)+(age85*.0157097)+
       (agef*.0107632)+(medicaid*-.0039739)+(3.589634);


*********************************************************************************
Calculate expected PCS same or better from the logit probablity.           *
Limit calculation to cases where baseline and follow-up PCS are non-missing.  *
*********************************************************************************;

if &C.pcs98 ne . and &R.pcs98 ne . then ExpPCSsb = exp(EPsb)/(1+exp(EPsb));


*************************************************************************************
*************************************************************************************
Step 2C: Calculate expected PCS better. Best possible (lowest lettered)          *
model is used for each case. First calculate EPb, the logit probablity of PCS    *
better.  Then calculate expected PCS better from the logit.                      *
*************************************************************************************
*************************************************************************************;


**********************************
PCS Better Model A               *
**********************************;

if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and pov2 ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EPb = (age*.004964)+(female*-.4577375)+(age75*.0021139)+(age85*-.010275)+
      (agef*.0069375)+(medicaid*.1440785)+(home*-.042036)+
      (pov2*.103442)+(hs*-.112051)+(married*.1275917)+
      (asian*.1383128)+(black*.1168021)+(hispanic*.0845965)+(-2.172204);


**********************************
PCS Better Model B               *
**********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EPb = (age*.0060665)+(female*-.5560831)+(age75*-.0014687)+(age85*-.003094)+
      (agef*.0080432)+(medicaid*.113725)+(home*-.0534817)+
      (hs*-.1346055)+(married*.0835041)+(asian*.1136843)+
      (black*.1529066)+(hispanic*.114252)+(-2.148534);


**********************************
PCS Better Model C               *
**********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne .
then

EPb = (age*.0070831)+(female*-.3755389)+(age75*-.0009189)+(age85*-.0090416)+
      (agef*.00546)+(medicaid*.1877218)+(-2.281332);


*********************************************************************************
```

```
Calculate expected PCS better from the logit probablity.                         *
Limit calculation to cases where baseline and follow-up PCS are non-missing. *
*********************************************************************************;

if &C.pcs98 ne . and &R.pcs98 ne . then ExpPCSb = exp(EPb)/(1+exp(EPb));


*************************************************************************************
*************************************************************************************
Step 2D: Calculate expected MCS same or better. Best possible (lowest lettered)    *
model is used for each case. First calculate EMsb, the logit probablity of MCS same *
or better.  Then calculate expected MCS same or better from the logit.              *
*************************************************************************************
*************************************************************************************;

**********************************
MCS Same+Better Model A          *
**********************************;

if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and pov2 ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EMsb = (age*-.0396762)+(female*-.3639759)+(age75*.0052427)+(age85*.0106093)+
        (agef*.0037154)+(medicaid*-.2678399)+(home*.1015457)+
        (pov2*-.2237441)+(hs*.2115211)+(married*-.2017257)+
        (asian*-.1026597)+(black*-.1326165)+(hispanic*-.078652)+(4.290668);


**********************************
MCS Same+Better Model B          *
**********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EMsb = (age*-.0413088)+(female*-.4101988)+(age75*.0053527)+(age85*.0150838)+
        (agef*.0043972)+(medicaid*-.2531204)+(home*.1099513)+
        (hs*.2636095)+(married*-.1268977)+
        (asian*-.0734853)+(black*-.1996325)+(hispanic*-.1342561)+(4.230617);


**********************************
MCS Same+Better Model C          *
**********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and hblack ne . and hasian ne . and hhisp ne .
then

EMsb = (age*-.0422551)+(female*-.5653867)+(hblack*-.2523381)+
          (hasian*.004378)+(hhisp*-.2387025)+(age75*.0032464)+
          (age85*.0142201)+(agef*.0069618)+(medicaid*-.3292756)+(4.469469);

*************************************************************************************
Calculate expected MCS same or better from the logit probablity.              *
Limit calculation to cases where baseline and follow-up MCS are non-missing. *
*************************************************************************************;

if &C.mcs98 ne . and &R.mcs98 ne . then ExpMCSsb = exp(EMsb)/(1+exp(EMsb));


*************************************************************************************
*************************************************************************************
Step 2E: Calculate expected MCS better. Best possible (lowest lettered)            *
model is used for each case. First calculate EMb, the logit probablity of MCS      *
better.  Then calculate expected MCS better from the logit.                        *
```

```
********************************************************************************
********************************************************************************;


*********************************
MCS Better Model A              *
*********************************;

if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and pov2 ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EMb = (age*.0096731)+(female*.8203816)+(age75*.0062538)+(age85*-.0018618)+
        (agef*-.0102308)+(medicaid*.148751)+(home*-.0326809)+
        (pov2*.210614)+(hs*-.2363332)+(married*-.0906781)+
        (asian*-.05392)+(black*.0251279)+(hispanic*.1929396)+(-2.385757);


*********************************
MCS Better Model B              *
*********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and home ne . and hs ne . and married ne . and
asian ne . and black ne . and hispanic ne .
then

EMb = (age*.0109386)+(female*.7174221)+(age75*.0086711)+(age85*-.0096264)+
        (agef*-.0084135)+(medicaid*.2036733)+(home*-.0314016)+
        (hs*-.270556)+(married*-.1512992)+
        (asian*-.0029232)+(black*.0629019)+(hispanic*.1814451)+(-2.319254);


*********************************
MCS Better Model C              *
*********************************;

else if age ne . and female ne . and age75 ne . and age85 ne . and agef ne . and
medicaid ne . and hblack ne . and hasian ne . and hhisp ne .
then

EMb = (age*.0156096)+(female*.6855663)+(hblack*.1435944)+
          (hasian*.0787052)+(hhisp*.2789385)+(age75*.0082538)+
          (age85*-.0025787)+(agef*-.0074519)+(medicaid*.324786)+(-2.986842);


********************************************************************************
Calculate expected MCS better from the logit probablity.                      *
Limit calculation to cases where baseline and follow-up MCS are non-missing. *
********************************************************************************;

if &C.mcs98 ne . and &R.mcs98 ne . then ExpMCSb = exp(EMb)/(1+exp(EMb));

***********************************************************************************
***********************************************************************************
Step 3: Calculate actual PCS and MCS beter, same, and same or better.            *
***********************************************************************************
***********************************************************************************
;

***********************************************************************************
Create variables for actual PCS same, actual PCS better, and actual PCS same or better *
***********************************************************************************;


************************************************************
Determine if follow-up survey was phone administration.    *
************************************************************;

&R.phon = substr(&R.srvdsp,1,1);
```

```
if &R.phon = 'T' then &R.tele = 1;
else &R.tele = 0;
drop &R.phon;

*********************************************************************************
Correct for effect of phone assessment by subtracting 1.9 from PCS scores if     *
phone survey.                                                                    *
********************************************************************************;


&C.pcs1 = &C.pcs98 - 1.9*tele1;
&R.pcs1 = &R.pcs98 - 1.9*&R.tele;


*********************************************************************************
Calculate difference score for PCS and determine if actual PCS score is better   *
between follow-up and baseline, or the same between follow-up and baseline.  Also *
create variable if PCS is better or the same.  PCS is better if follow-up minus   *
baseline is 5.66 or more.  PCS is the same if follow-up minus baseline is greater *
than -5.66 and less than 5.66.                                                   *
Limit calculation to cases with both a baseline and follow-up PCS.               *
********************************************************************************;


pcsdiff = &R.pcs1-&C.pcs1;

if pcsdiff ne . then do;

if pcsdiff ge 5.66 then actpcsb = 1;
else actpcsb = 0;

if -5.66 lt pcsdiff lt 5.66 then actpcssa = 1;
else actpcssa = 0;

if actpcsb = 1 or actpcssa = 1 then actpcssb = 1;
else actpcssb = 0;
end;

*********************************************************************************
Create variables for actual MCS same, actual MCS better, and actual MCS same or better *
********************************************************************************;

*********************************************************************************
Correct for effect of phone assessment by subtracting 4.5 from MCS scores if phone *
survey.                                                                          *
********************************************************************************;

&C.mcs1 = &C.mcs98 - 4.5*tele1;
&R.mcs1 = &R.mcs98 - 4.5*&R.tele;

*********************************************************************************
Calculate difference score for MCS and determine if actual MCS score is better between *
follow-up and baseline, or the same between follow-up and baseline.  Also create  *
variable if MCS is better or the same.  MCS is better if follow-up minus baseline is *
6.72 or more.  MCS is the same if follow-up minus baseline is greater than -6.72 and *
less than 6.72.  Limit calculation to cases with both a baseline and follow-up MCS.  *
********************************************************************************;

mcsdiff = &R.mcs1-&C.mcs1;

if mcsdiff ne . then do;

if mcsdiff ge 6.72 then actmcsb = 1;
else actmcsb = 0;

if -6.72 lt mcsdiff lt 6.72 then actmcssa = 1;
else actmcssa = 0;

if actmcsb = 1 or actmcssa = 1 then actmcssb = 1;
else actmcssb = 0;

end;
```

```
********************************************
Labeling model variables                 *
*********************************************;

Label

survdate = 'Survey date'
age      = 'Survdate - bdate/365.25'
deadcat  = '2-level dead in 2yr window after baseln'
dethvar  = '4-level dead/not dead variable'
age75    = 'Age greater than or = 75 Spline'
age85    = 'Age greater than or = 85 Spline'
female   = '1/0 variable using &C.gender'
hblack   = '1/0 variable for black using CMS race'
hasian   = '1/0 variable for asian using CMS race'
hhisp    = '1/0 variable for hispan using CMS race'
black    = '1/0 variable for black using &C.sv_rac'
asian    = '1/0 variable for asian using &C.sv_rac'
hispanic = '1/0 variable for hispan using &C.sv_rac'
agef     = 'Age/Female interaction'
medicaid = '1/0 variable for medicaid status'
tele1    = 'Telephone administered using &C.srvdsp'
pov2     = 'If income < $20,000'
hs       = 'High school graduate or GED'
married  = '1/0 variable for marital status'
home     = '1/0 variable for homeowner status'
highbp   = '1/0 variable for &C.highbp'
angnacad = '1/0 variable for &C.angcad'
chf      = '1/0 variable for &C.chf'
ami      = '1/0 variable for &C.ami'
othrhart = '1/0 variable for &C.othhrt'
stroke   = '1/0 variable for &C.stroke'
copd_etc = '1/0 variable for &C.copd_e'
gi_etc   = '1/0 variable for &C.gi_etc'
arthhip  = '1/0 variable for &C.athhip'
arthhand = '1/0 variable for &C.athhan'
sciatica = '1/0 variable for &C.sciatc'
diabetes = '1/0 variable for &C.diabet'
anycanc  = '1/0 variable for &C.anycan'
colncanc = '1/0 variable for &C.colnca'
brstcanc = '1/0 variable for &C.brstca'
proscanc = '1/0 variable for &C.prosca'
lungcanc = '1/0 variable for &C.lungca'
sci      = 'Disease grp with large pos. regr coeff'
mci      = 'Disease grp with moderate pos. regr coef'
uci      = 'Disease grp with near zero regr coeff'
nci      = 'Disease grp with negative regr coeff'
pfadl    = 'SF36 Phys Funct/Act Daily Living (0-100)'
pcsdiff  = '&C.pcs1 - &R.pcs1'
mcsdiff  = '&C.mcs1 - &R.pcs1'
&R.tele  = '1/0 variable for foll-up phone administr'
&C.pcs1  = '&C.pcs98 - 1.9*tele1'
&R.pcs1  = '&R.pcs98 - 1.9*&R.tele'
&C.mcs1  = '&C.mcs98 - 4.5*tele1'
&R.mcs1  = '&R.mcs98 - 4.5*&R.tele'
ExpDeath = 'expected prob. death'
actdeath = '1/0 actual death within two year window'
ExpPCSsb = 'expected PCS same+better'
ExpPCSb  = 'expected PCS better'
ExpMCSsb = 'expected MCS same+better'
ExpMCSb  = 'expected MCS better'
actpcssb = 'Actual PCS same+better'
actpcsb  = 'Actual PCS better'
actpcssa = 'Actual PCS same'
actmcssb = 'Actual MCS same+better'
actmcsb  = 'Actual MCS better'
actmcssa = 'Actual MCS same'
ED       = 'Logit probability of death'
EPsb     = 'Logit probability of PCS same+better'
```

```
EPb      = 'Logit probability of PCS better'
EMsb     = 'Logit probability of MCS same+better'
EMb      = 'Logit probability of MCS better'
&C.ssadod = 'SSA date of death';

length  age75 age85 agef 4 actdeath
female hblack hasian hhisp black asian hispanic
medicaid tele1 pov2 hs married home highbp angnacad chf
ami othrhart stroke copd_etc gi_etc arthhip arthhand sciatica
diabetes anycanc colncanc brstcanc proscanc lungcanc &R.tele
actpcsb actpcssa actpcssb actmcsb actmcssa actmcssb 3;

run;


********************************************************************
Person Level Analysis Stops and Plan Level Analysis Begins Here *
********************************************************************;

***********************************************************************************************
***********************************************************************************************
Step 4: Calculate system-level averages for actual death, actual change in PCS/MCS        *
***********************************************************************************************
***********************************************************************************************
;

Proc Summary nway data=&lib..&pers;
   Var actdeath actpcssb actpcsb actmcssb actmcsb;
   Output Out = Sysavge Mean = sysadeth sysapsb sysapb sysamsb sysamb;
run;

proc print data=Sysavge;
title 'System-level averages';
run;

***********************************************************************************************
***********************************************************************************************
Step 5: Calculate plan-level averages for actual and expected deaths, and for actual  *
and expected change in PCS/MCS (same or better, better).                              *
Also calculate N for plan-level death, PCS and MCS analysis.                          *
***********************************************************************************************
***********************************************************************************************;

Proc Summary nway data=&lib..&pers;
   Class &p.cntrnm;
   Var expdeath actdeath exppcssb actpcssb exppcsb actpcsb expmcssb actmcssb expmcsb
actmcsb;
   Output Out = Plnavge Mean =
       plnedeth plnadeth plnepsb plnapsb plnepb plnapb plnemsb plnamsb plnemb plnamb
       n = nd p1 np p2 p3 p4 nm;
run;



***********************************************************************************************
***********************************************************************************************
Step 6: Merge system and plan-level data and perform final calculations.        *
***********************************************************************************************
***********************************************************************************************;

**********************************************************
Create a link variable to merge system and plan-level data. *
**********************************************************;

Data sysavge;
  set sysavge;
Linkvar=1;
run;


Data plnavge;
```

```
   set plnavge;
drop p1-p4;
Linkvar = 1;
run;


****************************************************************
Merge system and plan-level data.                             *
****************************************************************;

Data &lib..&plan;
   Merge sysavge plnavge;
   By Linkvar;


****************************************************************
Create Plan-level deviation scores for:                       *
1=PCS Alive and Same or Better                                *
2=PCS Alive and Better                                        *
3=MCS Same or Better                                          *
4=MCS Better                                                  *
****************************************************************;


**********************************************************************************
Calculate system-level actual alive and PCS same/better, and alive and PCS better  *
**********************************************************************************;

sysapasb = (1-sysadeth)*sysapsb;
sysapab  = (1-sysadeth)*sysapb;


*************************************************************************************
Calculate plan-level actual and expected alive and PCS same/better,               *
alive and PCS better                                                              *
*************************************************************************************;

plnepasb = (1-plnedeth)*plnepsb;
plnapasb = (1-plnadeth)*plnapsb;
plnepab  = (1-plnedeth)*plnepb;
plnapab  = (1-plnadeth)*plnapb;


*************************************************************************************
Calculate plan-level deviation scores (actual minus expected) for PCS alive and same or *
better, PCS alive and better, MCS same or better, and MCS better                  *
*************************************************************************************;

pldpasb = plnapasb-plnepasb;
pldpab =  plnapab-plnepab;
pldmsb = plnamsb-plnemsb;
pldmb =  plnamb-plnemb;


*************************************************************************************
Calculate adjusted plan-level scores using logit transformation.                  *
Note: One plan has division by zero for lgtdmsb calculation. Set adjusted MCS same or  *
better to  be equal to actual MCS same/better for this plan.                      *
*************************************************************************************;

lgtdpasb = log(plnapasb/(1-plnapasb)) - log(plnepasb/(1-plnepasb));
lgtapasb = log(sysapasb/(1-sysapasb)) + lgtdpasb;
adjppasb = 1/(1+exp(-lgtapasb));

lgtdpab = log(plnapab/(1-plnapab)) - log(plnepab/(1-plnepab));
lgtapab = log(sysapab/(1-sysapab)) + lgtdpab;
adjppab = 1/(1+exp(-lgtapab));

lgtdmsb = log(plnamsb/(1-plnamsb)) - log(plnemsb/(1-plnemsb));
lgtamsb = log(sysamsb/(1-sysamsb)) + lgtdmsb;
```

```
adjpmsb = 1/(1+exp(-lgtamsb));
if adjpmsb=. then adjpmsb=plnamsb;

lgtdmb = log(plnamb/(1-plnamb)) - log(plnemb/(1-plnemb));
lgtamb = log(sysamb/(1-sysamb)) + lgtdmb;
adjpmb = 1/(1+exp(-lgtamb));


****************************************************************************
Using values for PCS alive & same or better, and PCS alive & better, calculate *
adjusted PCS better, PCS same, PCS worse.  Using values for MCS same or better *
and MCS better, calculate adjusted MCS better, MCS same, MCS worse.           *
****************************************************************************;

adjpcsb=adjppab;
adjpcss=adjppasb-adjppab;
adjpcsw=1-adjppasb;

adjmcsb=adjpmb;
adjmcss=adjpmsb-adjpmb;
adjmcsw=1-adjpmsb;


***********************************
Calculate deviation scores        *
***********************************;

dpsb = plnapasb - plnepasb;
dpb  = plnapab - plnepab;
dmsb = plnamsb - plnemsb;
dmb  = plnamb - plnemb;

***********************************
Calculate standard errors         *
***********************************;

sedpsb = sqrt ((((plnepsb)**2 * plnedeth * (1-plnedeth)) / nd)
                               +
        (((1-plnedeth)**2 * plnepsb * (1-plnepsb)) / np));

sedpb = sqrt(((((plnepb)**2 * plnedeth * (1-plnedeth)) / nd)
                               +
        (((1-plnedeth)**2 * plnepb * (1-plnepb)) / np));

sedmsb = sqrt ((plnemsb * (1-plnemsb)) / nm);
sedmb  = sqrt ((plnemb * (1-plnemb)) / nm);

***********************************
Calculate T statistics            *
***********************************;

tpsb = dpsb / sedpsb;
tpb  = dpb / sedpb;
tmsb = dmsb / sedmsb;
tmb  = dmb / sedmb;

*******************************************************
Labeling model variables                             *
*******************************************************;


Label

plnadeth = 'Plan actual death'
plnapb   = 'Plan actual PCS better'
plnapsb  = 'Plan actual PCS same+better'
plnedeth = 'Plan expected death'
plnepb   = 'Plan expected PCS better'
plnepsb  = 'Plan expected PCS same+better'
sysadeth = 'System actual death'
```

```
sysapb    = 'System actual PCS better'
sysapsb   = 'System actual PSC same+better'
ND        = 'Number for death analysis in plan'
NP        = 'Number for PCS analysis in plan'
NM        = 'Number for MCS analysis in plan'
sysapasb  = 'System actual PCS alive same+better'
sysapab   = 'System actual PCS alive better'
sysamsb   = 'System actual MCS same+better'
sysamb    = 'System actual MCS better'
plnepasb  = 'Plan expected PCS alive same+better'
plnapasb  = 'Plan actual PCS alive same+better'
plnepab   = 'Plan expected PCS alive better'
plnapab   = 'Plan actual PCS alive better'
plnemsb   = 'Plan expected MCS same+better'
plnamsb   = 'Plan actual MCS same+better'
plnemb    = 'Plan expected MCS better'
plnamb    = 'Plan actual MCS better'
pldpasb   = 'Plan deviation PCS alive same+better'
pldpab    = 'Plan deviation PCS alive better'
pldmsb    = 'Plan deviation MCS same+better'
pldmb     = 'Plan deviation MCS better'
lgtdpasb  = 'Logit Act.-Pred. PCS alive same+better'
lgtapasb  = 'Logit plan Dev. PCS alive same better'
adjppasb  = 'Adjusted PCS alive same+better'
lgtdpab   = 'Logit Act.-Pred. PCS alive better'
lgtapab   = 'Logit plan dev. PCS alive better'
adjppab   = 'Adjusted PCS alive better'
lgtdmsb   = 'Logit Act.-Pred. MCS same+better'
lgtamsb   = 'Logit plan dev. MCS same+better'
adjpmsb   = 'Adjusted MCS same+better'
lgtdmb    = 'Logit Act.-Pred. MCS better'
lgtamb    = 'Logit plan dev. MCS better'
adjpmb    = 'Adjusted MCS better'
linkvar   = 'Variable to link person/plan data'
adjpcsb   = 'Adjusted PCS better'
adjpcss   = 'Adjusted PCS same'
adjpcsw   = 'Adjusted PCS worse'
adjmcsb   = 'Adjusted MCS better'
adjmcss   = 'Adjusted MCS same'
adjmcsw   = 'Adjusted MCS worse'
dpsb      = 'Plan deviation PCS same+better'
dpb       = 'Plan deviation PCS better'
dmsb      = 'Plan deviation MCS same+better'
dmb       = 'Plan deviation MCS better'
sedpsb    = 'Standard error Dev. PCS same+better'
sedpb     = 'Standard error Dev. PCS better'
sedmsb    = 'Standard error Dev. MCS same+better'
sedmb     = 'Standard error Dev. MCS better'
tpsb      = 'T statistic - PCS same+better'
tpb       = 'T statistic - PCS better'
tmsb      = 'T statistic - MCS same+better'
tmb       = 'T statistic - MCS better';


run;


Proc Means;
   var plnapasb plnepasb plnapab plnepab plnamsb plnemsb plnamb plnemb
       pldpasb pldpab pldmsb pldmb lgtdpasb lgtapasb adjppasb
       lgtdpab lgtapab adjppab lgtdmsb lgtamsb adjpmsb lgtdmb lgtamb adjpmb
       adjpcsb adjpcss adjpcsw adjmcsb adjmcss adjmcsw dpsb dpb dmsb dmb
       sedpsb sedpb sedmsb sedmb tpsb tpb tmsb tmb;
title 'Summary across all plan-level results';
run;


*****************************************************************
Step 7:  Calculate overall F statistics for plan comparisons   *
            for Death, PCS same+better, PCS better,             *
```

```
            MCS same+better, MCS better.                        *
*****************************************************************;

data &lib..&pers2;
   set &lib..&pers;

*****************************************************************
Step 7A: Calculate difference scores for Death, PCS same+     *
        better, PCS better, MCS same+better, and MCS better   *
        by subtracting the Expected value from the Actual     *
        value for each measure.                               *
*****************************************************************;

resdeath = ActDeath- ExpDeath;
resPCSb  = ActPCSb - ExpPCSb;
resPCSsb = ActPCSsb- ExpPCSsb;
resMCSb  = ActMCSb - ExpMCSb;
resMCSsb = ActMCSsb- ExpMCSsb;

*****************************************************************
Step 7B: Calculate overall F statistics using ANOVA, where    *
        reporting unit is the independent variable and each   *
        difference score is the dependent variable            *
*****************************************************************;

proc glm;
   class &p.cntrnm;
   model resdeath = &p.cntrnm;
   Title 'Overall F Test for Death';
run;

proc glm;
   class &p.cntrnm;
   model resPCSsb = &p.cntrnm;
   Title 'Overall F Test for PCS Same+Better';
run;

proc glm;
   class &p.cntrnm;
   model resPCSb = &p.cntrnm;
   Title 'Overall F Test for PCS Better';
run;

proc glm;
   class &p.cntrnm;
   model resMCSsb = &p.cntrnm;
   Title 'Overall F Test for MCS Same+Better';
run;

proc glm;
   class &p.cntrnm;
   model resMCSb = &p.cntrnm;
   Title 'Overall F Test for MCS Better';
run;
quit;


*****************************************************************
Step 7C: Identify Outliers (if warranted based on F test)     *
*****************************************************************;

proc print data = &lib..&plan;
      var &p.cntrnm tpsb;
      where tpsb ge 2;
      title 'PCS Better than Expected';
run;

proc print data = &lib..&plan;
      var &p.cntrnm tpsb;
      where tpsb le -2;
```

```
            title 'PCS Worse than Expected';
run;

proc print data = &lib..&plan;
      var &p.cntrnm tmsb;
      where tmsb ge 2;
      title 'MCS Better than Expected';
run;

proc print data = &lib..&plan;
      var &p.cntrnm tmsb;
      where tmsb le -2;
      title 'MCS Worse than Expected';
run;
proc print data = &lib..&plan label;
      var &p.cntrnm tpsb tmsb;
      where tpsb ge 2;
      title 'PCS Better than Expected';
run;

proc print data = &lib..&plan label;
      var &p.cntrnm tpsb tmsb;
      where tpsb le -2;
      title 'PCS Worse than Expected';
run;

proc print data = &lib..&plan label;
      var &p.cntrnm tmsb tpsb;
      where tmsb ge 2;
      title 'MCS Better than Expected';
run;

proc print data = &lib..&plan label;
      var &p.cntrnm tmsb tpsb;
      where tmsb le -2;
      title 'MCS Worse than Expected';
run;

proc freq data = &lib..&pers;
      table actdeath;
      title 'Actual Death';
run;

data test1;
      set &lib..&plan;
      npcsb=100*sysapab;
      npcss=100*(sysapasb-sysapab);
      npcsw=100*(1-sysapasb);
      nmcsb=100*sysamb;
      nmcss=100*(sysamsb-sysamb);
      nmcsw=100*(1-sysamsb);
      label npcsb="National PCS Better"
      npcss="National PCS Same"
      npcsw="National PCS Worse"
      nmcsb="National MCS Better"
      nmcss="National MCS Same"
      nmcsw="National MCS Worse"
      run;
proc freq; tables npcsb npcss npcsw nmcsb nmcss nmcsw;
      title "Values for &C PM Report";
      run;
```