IMPLEMENTING THE HEDIS®
MEDICARE HEALTH OUTCOMES SURVEY

Medicare Health Outcomes Survey: An Alternative Case-Mix Methodology

Prepared By

Alfredo J. Selim MD, MPH [1,2,3,4]
Sheikh Usman Iqbal MD, MPH[2]
William Rogers Ph.D[1,5]
Shirley X. Qian M.S[1]
Benjamin G Fincke, MD[1,2,4]
Jim Rothendler MD[1,2,4]
Lewis E. Kazis Sc.D.[1,2]

1 Center for Health Quality, Outcomes, and Economic Research, A Health Services Research and Development Field Program, VA Medical Center, Bedford, Massachusetts.
2 Center for the Assessment of Pharmaceutical Practices (CAPP), Department of Health Policy and Management, Boston University School of Public Health, Boston MA
3 Section of General Internal Medicine and Emergency Services, Boston VA Health Care System, West Roxbury, Massachusetts
4 Boston University School of Medicine, Boston, Massachusetts
5 Senior Research Scientist, The Health Institute, New England Medical Center

Prepared under subcontract to the National Committee for Quality Assurance for the Centers for Medicare and Medicaid Services under Contract Number HHSM-500-2004-00015I-0001
OY1 Deliverables 405-407
August 22, 2007

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.

IMPLEMENTING THE HEDIS®
MEDICARE HEALTH OUTCOMES SURVEY

Medicare Health Outcomes Survey: An Alternative Case-Mix Methodology

TABLE OF CONTENTS

1.0 EXECUTIVE SUMMARY

1.1 Background and Purpose

The Medicare Health Outcome Study (HOS) provides longitudinal health status data on Medicare patients in managed care plans for use in quality improvement activities, plan accountability, public reporting, and improving health. The Medicare HOS program uses the outcomes of change in health status after adjustment with a fairly complex multi-modeling case-mix methodology. The purpose of this project is to use a theory and evidence-based hierarchical approach to develop and test an alternative case-mix adjustment methodology that is simpler and more parsimonious.

1.2 Methods

The data analyzed in this project were obtained from the Medicare HOS cohort 7 (2004- 2006) at baseline and follow-up 2 years later. We developed a case-mix adjustment model using a "hierarchical" approach to select a core set of case-mix domains. The development of the model used a sequential core set of outcome-specific case-mix domains for the outcomes of mortality, Physical Component Summary (PCS) and Mental Component Summary (MCS) the same or better. We then compared the performance measurement results between the new case-mix methodology and the current HOS case-mix methodology used.

1.3 Results

The new case-mix adjusted models performed reasonably well in cross-validated tests of discrimination (c-statistics = 0.80, 0.67 and 0.69 for mortality, PCS the same or better and MCS the same and better, respectively) and calibration. The adjusted plan rates with the new case-mix methodology correlated significantly with those of the HOS case-mix methodology (r=0.91, p<.0001 for being alive with PCS the same or better and r=0.89, p<.0001 for MCS the same or better). However, the new case-mix methodology detected fewer positive plan outliers than the HOS case-mix method (1 vs. 6).

1.4 Conclusions

Our results suggest that the adjusted proportion of patients with change in health status and the plan rankings are similar regardless of whether the current HOS method or alternative case-mix methodology is used. However, the current HOS case-mix methodology detects a greater number of positive plan outliers. Further studies should examine what differences in care structures and processes contribute to better outcomes in those high-performing plans.

1.5 Recommendations for CMS

On the basis of the findings of this report we recommend the continued use of the current HOS case-mix methodology because of its conceptual relevance and superior ability to detect more completely high-performing plans.

2.0 INTRODUCTION

The Medicare Health Outcomes Survey (HOS) is an annual evaluation of the physical and mental health of Medicare beneficiaries enrolled in managed care plans.[1] Its development was the result, in part, of the proliferation of Medicare managed care plans in the 1990's, and the need for the Centers for Medicare & Medicaid Services (CMS) to monitor the health outcomes of Medicare beneficiaries.

For reports on the performance of health care plans to be effective, profiling must be done using the best statistical methods to control for variations in case-mix characteristics.[2] Case-mix refers to those patient specific characteristics that are outside the control of the plan and that could contribute to better or worse outcomes, but over which the plan has little or no influence. Commonly used case-mix adjustment methodologies often contain some of the following deficiencies.[3] They rely on small data sets that pose limitations in the completeness of case-mix data. They use approaches that may not be applicable to the majority of the patients or they use statistical standards that may not be adequate to handle patients with incomplete or missing case-mix data.

In a technical report commissioned by CMS, RTI International compared expected versus actual follow-up health status for profiling health care systems serving Fee-for-Services Medicare beneficiaries.[4] They developed statistical models to adjust the predicted two-year outcomes for the case-mix characteristics of the groups under study. They used a limited number of case-mix variables including age, baseline physical and mental health status in addition to congestive heart failure and angina/coronary artery disease for predicting physical health at follow-up and depression for predicting mental health at follow-up. They imputed the follow-up health scores for respondents who died between baseline and follow-up. However, they found that their approach to case-mix adjustment had limited value in identifying differences among health care systems that have clinical or policy relevance. There are several possible explanations. First, the limited case-mix models may have produced inaccurate expected levels for comparisons. Second, exclusion of patients with incomplete or missing data may have biased the results.

The HOS performance measurement results are computed using a rigorous case-mix methodology that was developed by Rogers et al.[5] This approach relies on data obtained from beneficiaries, who provide information on their baseline health status (from the SF-36® Health Survey and summarized as physical health (PCS) and mental health (MCS)), chronic conditions, and sociodemographic data. Multivariate logistic regression models are used to develop a series of eight different death models and three different models for the outcomes of the same or better physical health (or mental health) as measured by the SF-36® Health Survey over 2 years.[5] The models are theory-driven, and range from the more complex (e.g., controlling for a number of sociodemographic characteristics, specific chronic conditions, and specific health status domains) to the more basic (e.g., controlling for age, gender, race, and Medicaid status only). Expected values for each beneficiary are calculated using the most complex model possible, given the beneficiary's pattern of missing data. Rogers et al. found that this method was able to identify high- and low-performing Medicare managed care plans.[6] One of the major missions of the HOS is to provide an ongoing evaluation of the health plans for purposes of public accountability. This mission is quite public and results in the dissemination of outcomes to a wide-ranging audience from providers and administrators to policy makers. Consequently the case-mix methodology needs to be fairly straightforward and understandable to both a technical and lay audience. In this environment a simpler and more parsimonious case-mix methodology may be more adequate.

Thus, the purpose of this project was to develop and test an alternative case-mix adjustment methodology for comparisons of managed care plans. Specifically, this project used a “hierarchical” approach to select a core set of case-mix domains that theory and prior research suggest are important determinants of physical and mental health. Potential advantages of this approach include a greater degree of simplicity, understandability, conceptual meaningfulness, and parsimony in data elements. We then compared the performance measurement results between the new case-mix methodology and the current HOS case-mix methodology used.

3.0 METHODS

3.1 Study Population

We used the Medicare HOS cohort 7 (2004-2006) since it is considered the "transition cohort" and provides a data set with baseline SF-36® Health Survey and the Veterans RAND 12-Item Health Survey (VR-12) at follow-up that can be used for evaluating different methods of case-mix adjustment (Figure 1).[6,7,8] Beneficiaries were randomly sampled from each managed care plan participating in the Medicare Advantage Program in 2004 for the cohort 7 baseline administration. There were 161 contract/market areas, and 1,000 beneficiaries were randomly sampled from each (if a contract/market area had fewer than 1,000 members, all beneficiaries were included). The total number of beneficiaries sampled was 159,311, including both the aged and disabled. By the time of the two-year follow-up, some plans no longer offered managed care to Medicare beneficiaries; these organizations were not included in the follow-up survey administration. There also was a consolidation of contract numbers within some health plans and a reclassification of market areas within some states during this time period, resulting in a total of 150 plans at the end of two years.

The analysis was limited to beneficiaries age 65 years and older (N=147,955). Of these individuals, 98,890 (66.8%) completed a baseline HOS survey (complete, partial complete, or break-off) and had enough data to score the physical (PCS) or mental (MCS) component summary scores using a previously validated Modified Regression Estimation (MRE) algorithm

---

® MOS SF-12 and SF-36 are registered trademarks of the Medical Outcomes Trust.

for missing data.[9] The MRE algorithm has been validated against other approaches for imputing responses to the SF-36® and VR-12 in ambulatory patient populations where regression towards the mean is a concern with data that include missing values used to calculate PCS and MCS scores.[10] With the use of the MRE algorithm, PCS and MCS can be computed in as many as 90% of the cases in which one or more survey responses are missing. Among the 98,890 Medicare beneficiaries, 7,801 died during the 2-years of follow-up. Of the 91,089 beneficiaries who were alive during the 2-years of follow-up, 5,452 were not resurveyed because their plan no longer participated in Medicare managed care (involuntary disenrollment). Furthermore, 16,966 were not resurveyed because they were no longer enrolled in the plan in 2006 (voluntary disenrollment). Of the 68,671 beneficiaries who were surveyed at 2 years, 54,741 completed a baseline HOS survey (complete, partial complete, or break-off) and had enough VR-12 data to score the physical (PCS) or mental (MCS) component summary scores using the MRE algorithms.

3.2 Outcome Measures

The HOS used the SF-36® Health Survey at baseline and the VR-12 was used for follow-up. The VR-12 is a modification of the RAND 36 Item Health Surveys 1.0 that was developed at the RAND Corporation as part of the Medical Outcomes Study. It differs from the SF-36® Health Survey in the use of 5-point response choices for the role limitations due to physical problems and the role limitations due to emotional problems.[11] We summarized the health surveys into physical (PCS) and mental component (MCS) scales by applying a linear t-score transformation with a mean of 50 and a standard deviation of 10 based upon the 1990 U.S. population norms.[12,13]

For the analysis, the PCS and MCS scores at baseline were calculated based upon 12 items of the SF-36® Health Survey, equivalent to the MOS SF-12. The follow-up PCS and MCS scores were calculated using the VR-12 conversion formulas equivalent to the MOS SF-12.[11] The HOS used two modes of health survey administration: mail or telephone. We adjusted for the effect of telephone administration by subtracting 1.9 points from PCS scores and 4.5 points from MCS scores based on a previous validated approach.[14] This was done for both the baseline and follow-up scores.

We used CMS data to ascertain vital status. A person is defined as dead at follow-up if he or she died within two years of the baseline survey date. 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 having died and were included in the death analysis.

The primary outcomes examined were two: (1) “being alive with PCS the same or better” (vs. PCS worse or death) and (2) “MCS the same or better” (vs. MCS worse). They were designated as the primary outcomes of interest because health maintenance, rather than improvement, is a realistic clinical goal for many seniors. The cut-off points for the operational definition of “the same or better” than would be expected by chance were ‘two standard errors of the measurement’ (SEM) for a single score or 1.414 standard errors of change. From this definition, the categories of change in PCS were classified as: 1) “the same” (or unchanged) between -5.66 points and +5.66 points; 2) “better” as greater than +5.66 points; and 3) “worse” as lower than -5.66 points. The categories of change in MCS were classified as: 1) “the same” (or unchanged) between -6.72 points and +6.72 points; 2) “better” as greater than +6.72 points; and 3) “worse” as lower than -6.72 points. The primary outcomes were expressed in terms of the percentage of beneficiaries who were either "alive with PCS the same or better" or "MCS the same or better" at 2-years follow-up. For MCS, death was not included in the metric.

3.3 Case-Mix Variables

Based on prior work,[15,16] we identified three domains of case-mix characteristics: sociodemographics, comorbidities and baseline health status (Table 1).

Since the risk for health outcomes differs by demographic subgroup, as has been previously theorized and empirically shown using the VR-12, we have selected the following sociodemographic variables: age, race/ethnicity, marital status, education, income and Medicaid eligibility.[8,9,17] Age of the beneficiary was calculated using the birth date and the survey date, in years. We grouped beneficiaries into 4 racial/ethnic groups: Whites, African-Americans, Hispanics and others. The marital status was dichotomized as being currently married and not married. The education and income were also dichotomized, at least a high school education (12 or more years) vs. less than 12 years of education and income less than $20,000 vs. equal or higher $20,000, respectively. The current HOS case-mix methodology uses similar case-mix variables with the addition of age and gender interaction term and homeowner status for the majority of its 14 models (Table 1 and Appendix A).

Because substantial variation in health status among patients with different diagnoses exists, we used a group of conditions that are commonly encountered in clinic visits and are known to be major indicators of health status.[18] The self-reported diagnoses were acute myocardial infarction (AMI), coronary artery disease, congestive heart failure, other heart conditions, stroke, hypertension, diabetes, chronic obstructive pulmonary disease (COPD), asthma, cancer (other than skin cancer), gastrointestinal disorder, arthritis (hip and hand), and sciatica. They were scored 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. We also used four indicators of current cancer treatment, including breast, prostate, colon and lung cancer. The four items on treatment for specific types of cancer are scored 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. The current HOS case-mix methodology uses co-morbidity information only for the mortality models.

In the domain of baseline health status, we used the baseline PCS and MCS scores because they are important predictors of health.[19] The baseline PCS and MCS scores were included as continuous variables. This was because there are no universally accepted cut-off levels for PCS and MCS scores and the correlations of PCS and MCS scores with sociodemographic variables is quite linear. This has been confirmed in previous work through scatterplot designs.[20] The current HOS case-mix methodology uses baseline health information only for the mortality models. The models include SF-36 Physical Functioning/Activities of Daily Living Index, SF-36 General Health scale, SF-36 Social Functioning scale, one-item measure of General Health compared to others and baseline PCS and MCS scores (linear).

All characteristics of individual patients were obtained in their baseline surveys, except for Medicaid eligibility that was reported by the states.

3.4 Statistical Analysis

The data analysis included: (1) case-mix model development; (2) calculation of expected outcomes for each beneficiary; (3) calculation of plan-level results and (4) comparison between the new case-mix approach and the current HOS case-mix methodology for assessing plan performance.

3.5 Case-Mix Model Development

All beneficiaries age 65 and older who completed the HOS at baseline and had baseline PCS or MCS scores were included in developing case-mix models for mortality (including beneficiaries whose plans subsequently left Medicare managed care). Beneficiaries age 65 years and older who completed the HOS survey at baseline and follow-up and for whom PCS and/or MCS could be calculated at both time points were included in the development of the PCS and MCS case-mix models.

We used a random sample of 2/3 of the study population (derivation sample) to develop the models. This was done in several stages to examine how the selection of the domains of sociodemographics, comorbidities and baseline health status affects the performance of the models. We used two measures to assess the performance of the models. First, we calculated the c-statistic, which reflects the predictive power of the models to discriminate among patients by ordering them according to rates of the outcome event. A c-statistic value of less than 0.5 indicates poor discriminatory power of the model. Second, we used the Hosmer-Lemeshow statistic test to evaluate the calibration of the model. Patients were divided into deciles, based on the expected risk for decline in PCS or MCS. Within each decile, the expected rate was compared with the observed rate. A p-value greater than 0.05 indicates a good fit. We then applied the regression coefficients from the derivation models onto the remaining 1/3 of the sample (validation sample). We also calculated pseudo R-squares, which are measures of the improvement in fit of the regression model that is due to the independent variables.

Calculation of Expected Outcomes for Each Beneficiary

Expected probability values were calculated for three outcomes: death; PCS the same or better; and MCS the same or better. Logistic regression was used to adjust for case-mix and to calculate expected outcomes for each beneficiary. The parameter estimates from the death regression model were used to calculate the logit probability of death for each beneficiary. This logit probability was then transformed into an expected probability of death, using a standard transformation:

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

where e^{ED} is the logit of Expected Death

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

For example, the expected probability of PCS the same or better is:

$$\text{Expected PCS the same or better} = e^{EPsb}/(1+e^{EPsb})$$

where e^{EPsb} is the logit of Expected PCS the same or better

There were important methodological differences between the new case-mix method and the currently used HOS case-mix method used in the calculation of the expected values. First, the new case-mix methodology used a single model to calculate the expected probability values for three outcomes: death, PCS the same or better, and MCS the same or better. In contrast, the current HOS case-mix methodology uses a series of eight different death models, three different PCS models, and three different MCS models (Appendix A).

Second, the new case-mix methodology used a correction factor to account for those beneficiaries with missing case-mix information. The correction factor was calculated as the difference between the observed value of those beneficiaries with case-mix data and those with missing case-mix data and multiplied by the proportion of beneficiaries with missing case-mix data. In contrast, the current HOS case-mix methodology calculates the expected values for each beneficiary using the most complex model possible given the beneficiary's pattern of missing data. If a beneficiary has all needed independent variables for the most comprehensive model, then their expected score is calculated using that model. If not, then the next less complex model is used if all needed independent variables for that model were available, and so on. One model is used for each beneficiary for each outcome, given the beneficiary's pattern of missing data. Details about the HOS case-mix methodology used are provided in the technical report submitted to CMS in June 2004.[5]

Calculation of Plan-Level Results

Health plans were evaluated on the percentage of beneficiaries whose health status was maintained or improved over two years. The outcomes were two: 1) being alive with PCS the same or better at 2-years follow-up and 2) MCS the same or better at 2-years follow-up.

The calculation of the overall plan-level results was done in several steps and followed the current HOS performance measurement process. First, as discussed above, a variable was created to indicate if each beneficiary in a plan who completed the baseline survey actually died during the two-year follow-up period. Second, for those beneficiaries who completed both the baseline and follow-up surveys, a variable was created to indicate if the PCS score was better or not at the two-year follow-up period. Third, an expected death rate was calculated for each beneficiary within a plan using the case mix models. Fourth, an expected PCS the same or better rate was calculated for each beneficiary using logistic regression techniques (detailed above). Neither of these calculations up to this point includes a variable for plan. To summarize data for all beneficiaries within a plan, the mean expected death rate (E_d) was calculated for all beneficiaries in the plan, along with the mean expected "PCS the same or better" rate (E_{pbs}). The expected "being alive with PCS the same or better" for the plan was $(1-E_d)*E_{pbs}$.

For the same beneficiaries within the plan, the mean observed (unadjusted) death rate (O_d) and mean observed "PCS the same or better" rate (O_{pbs}) were calculated across all beneficiaries. The unadjusted "being alive with PCS the same or better" rate for the plan was $(1-O_d)*O_{pbs}$.

An adjusted plan percent "being alive with PCS the same or better" was also calculated by combining the overall (national) results and the plan deviation score, using a logit transformation. The overall formulas for plan results were based on subtraction (observed minus expected) in terms of probabilities, but the standard errors are based on estimation in logistic terms. Therefore, standard errors, t-statistics, and confidence intervals were calculated in logistic terms and transformed back to probabilities for presentation.

Comparison Between the New Case-Mix Approach and the Current HOS Case-Mix Methodology for Assessing Plan Performance

We first examined the level of agreement between these two case-mix methodologies regarding the profiling of managed care plans. First, we plotted the plan rates with the new case-mix methodology versus the current HOS case-mix methodology and calculated their associations using a simple Pearson product moment correlation and associated P-value. We then examined the level of agreement between these two case-mix approaches regarding the identifications of plan outliers. We determined whether plans differed significantly on the two main outcome measures and reported the overall F statistic. If the overall F test was significant, then a *t* statistic was used to express the significance of each plan difference from the overall national results. The *t* statistics were calculated by dividing the plan deviation for the outcome ("being alive with PCS the same or better" and "MCS the same or better") by the standard error of the deviation. Plans that had a t statistic $\geq$ 2 were designated as significantly better than expected on that outcome (positive outliers), while plans that had a t statistic $\leq$ -2 were designated as significantly worse than expected (negative outliers), compared to the overall national results. This approach for identifying outlier type has traditionally been used by CMS in reports to the plans.

4.0 RESULTS

Table 2 shows the socio-demographic characteristics of the Medicare cohort 7. The sample size consisted of 147,955 Medicare enrollees. The study participants had a mean age of 76.3 years (SD±6) with representation of 40.8% males, 85.1% white, 54.9% married and 29.7% less than high school education. Forty seven percent of the participants had income of less than $20,000, 72.1% owned their homes, and 7.4% were Medicaid eligible. The average number of co-morbid conditions was 2.8 (SD± 2). Hypertension had the highest representation (62%), followed by any cardiac conditions (59%), arthritis of the hip (44.9%) and hand (38.5%), sciatica (23%), and diabetes (20.1%). A small proportion of the population (< 3%) was under treatment for breast, prostate, lung or colon cancer. The mean baseline PCS and MCS scores were 39.6 (SD±12) and 51.9 (SD±10), respectively. The PCS scores for our sample are substantially lower by close to one standard deviation below the norm of the general U.S. population that has a mean score of 50. MCS is slightly higher by about 0.2 of one standard deviation above the norm for the general U.S. population.

Table 3 shows the hierarchical multidimensional case-mix adjustment models for mortality, PCS the same or better and MCS the same or better in the derivation sample. The performance of the models improved with the block wise addition of different domains of case-mix. The models with only sociodemographic variables gave c-statistic values of 0.71, 0.53, and 0.57 for mortality, PCS the same or better and MCS the same or better, respectively. The addition of comorbidities into the model slightly improved the c-statistic values. The models with all three case-mix domains (sociodemographics, comorbidities and baseline PCS and MCS scores) increased the c-statistic values to 0.80, 0.67, and 0.69 for mortality, PCS the same or better and

MCS the same or better, respectively. When compared with the current HOS case-mix methodology, the c-statistic values were similar (Appendix A). The calibration of the new models also improved with the sequential addition of different dimensions of case-mix as shown by the Hosmer-Lemeshow tests (for mortality ($[chi]^2$ = 3.447 and P = 0.903), PCS the same or better ($[chi]^2$ = 10.814 and P = 0.212), and MCS the same or better ($[chi]^2$ = 8.191 and P = 0.415)). The pseudo R-square values ranged from 0.080 to 0.178 for mortality, from 0.002 to 0.063 for PCS the same or better and from 0.009 to 0.076 for MCS the same or better. They were similar to those pseudo R-square values in the current HOS case-mix methodology. The performance of the new models was confirmed in the validation sample (results not shown).

Table 4 summarizes the associations of individual patient characteristics and mortality, PCS the same or better and MCS the same or better. The case-mix variables that were more likely associated with dying were: old age; male; diabetes; stroke; congestive heart failure; COPD/asthma; cancer; treatment for breast cancer and lower baseline PCS and MCS scores. In contrast, those patients with hypertension, arthritis, sciatica or having treatment for prostate cancer were less likely to die. The individual patient characteristics that were more likely associated with having PCS the same or better were young age, income higher than $20,000 and high baseline MCS scores. Those with high PCS at baseline and those with diabetes, hypertension, heart disease, stroke, COPD/asthma, arthritis or sciatica were less likely to have PCS the same or better as the outcome. The individual patient characteristics that were more likely associated with having MCS the same or better were young age, not married, education higher than 12 years, income higher than $20,000, and high baseline PCS scores. Medicaid eligible patients, Asians, those with stroke and those with high MCS at baseline were less likely to have MCS the same or better as the outcome.

Figure 2 and 3 show the scatter plot diagrams for comparing the rates of 150 Medicare managed care plans with the new case-mix methodology versus the current HOS case-mix methodology. The plan rates with the new case-mix method were similar to those with the HOS case-mix method as shown by the proximity of the case-mix adjusted values to the diagonal of the figures (denoting no difference). This was confirmed by the significant correlation coefficients between the plan rates with the new case-mix methodology and the current HOS case-mix methodology ($r=0.91$, $p<.0001$ for being alive with PCS the same or better and $r=0.89$, $p<.0001$ for MCS the same or better, respectively).

Table 5 shows the comparisons of the individual plan outliers between the new case-mix methodology and the current HOS case-mix methodology. The new case-mix methodology detected fewer positive plan outliers when compared to the HOS case-mix methodology. The new case-mix methodology detected only one positive plan outlier with the adjusted outcome of being alive with PCS the same or better when compared to the current HOS case-mix methodology that identified five additional positive plan outliers. The new case-mix methodology and the current HOS case-mix methodology detected a similar number of negative plan outliers (8 vs.7 respectively). Regarding the adjusted outcome of MCS the same or better, both methods detected the same 4 positive plan outliers, however, the current HOS case-mix method detected an extra positive plan outlier. Both methods also detected the same 4 negative outlier plans but the new case-mix methodology detected one extra negative plan outlier.

5.0 DISCUSSION

Accurate information on health outcomes has become an expectation of regulatory and accreditation agencies.[21] Important decisions, such as reimbursements and accreditations, will be based on perceived performance. Our study showed that it is feasible to develop clinically credible case-mix adjustment models with good statistical properties for the outcomes of being alive with PCS the same or better and MCS the same or better using a theory and evidence-based hierarchical approach. The resulting models produced case-mix adjusted rates at the plan level, which were similar to those from the current HOS case-mix methodology.

Using the two different case-mix methodologies, we found plans that were significantly better than expected on the outcome of change in health status (positive outliers) suggesting that some plans are high-performers. This opens the possibility of examining these plans to identify processes of care or management practices that may serve as models of best practices. In contrast, plan rates that are significantly worse than expected are often attributed to poor quality of care and may serve to identify those plans that need to improve health care services through activities such as disease management programs or behavioral health practice guidelines.

Our new case-mix adjusted model performed well in cross-validated tests of discrimination and calibration. The associations among socio-demographic characteristics, diagnoses, baseline health status and the outcome of change in health status were consistent with the literature.[22,23] Although the models were developed in a heterogeneous ambulatory population, their c-statistic values were equal or superior to values that have been obtained in inpatient populations.[24,25] Tierney et al developed a case-mix adjusted mortality model of primary care patients with congestive heart failure using electronic medical records that resulted in a c-statistic of 0.76.[26] In a previous work, we also obtained similar c-statistic values in a population receiving ambulatory care in the Veterans Health Administration.[27] Finally, the model performances between the new case-mix methodology and the current HOS methodology were similar.

There are, however, several important differences between the new case-mix methodology and the current HOS case-mix methodology. First, the new case-mix methodology included baseline PCS and MCS scores in models for PCS (and MCS) the same or better. This is controversial because their coefficients in the model are influenced by the baseline score measurement error and inter-temporal correlation. However, other researchers have noted that high baseline scores meaning better functioning are predictors of worsening change in functional status and low baseline scores are predictors of improvement in functional status.[28,29] Appendix B shows that the new case-mix models with baseline PCS and MCS scores detects more low-performing plans than those without baseline PCS and MCS scores.

Second, the new case-mix methodology detects fewer high-performing plans (positive outliers) than the current HOS case-mix methodology used by CMS. This can be explained in part by the differences in handling missing case-mix information by these methodologies. The new case-mix methodology uses a correction factor to account for missing case-mix information. The correction factor was calculated as the difference between the observed value of those beneficiaries with case-mix data and those with missing case-mix data and multiplied by the proportion of beneficiaries with missing case-mix data at the plan level. In contrast, the HOS case-mix methodology uses a series of eight different death models, three different PCS models, and three different MCS models, because not all beneficiaries have data for all of the independent variables that could be used to calculate an expected score. Since models with fewer case-mix adjusters can be used given the beneficiary's pattern of missing data, this might result in expected values that are lower than those from the new case-mix methodology. As a result the bar for the expected is somewhat lower and plans are more likely to fall within the boundaries that define positive outliers.

Third, the new case-mix methodology and the current HOS case-mix methodology detect a few different low-performing plans (negative outliers). This can be explained in part by the calculation of the correction factor that was done at the plan level. This might leave the bar for the expected at a higher level with the new case-mix methodology. In contrast, the HOS case-mix methodology handles incomplete case-mix information at the patient level. This might correct the bar for the expected to slightly lower and those same plans would not be considered outliers.

We should note several limitations of this study. The validity of our case-mix adjustment methodology is based on the accuracy and completeness of the specific covariates included in the adjustment models, which were patient self-reported. Controlling for socio-demographics, co-morbidities and baseline health status explained only a fraction of the variance in the outcomes measured as reflected by the low pseudo R-square values in both case-mix methodologies. The same has been true in other studies.[30] This represents an opportunity for future work that may lead to improved model performance and possibly further change in judgments of plan performance.

In summary, the results show that it is feasible to develop clinically credible risk adjustment models with good statistical properties for the outcomes of change in health status using a hierarchical approach. The adjusted proportion of patients with change in health status and the plan rankings are similar regardless of whether the current HOS or the alternative model is used to case-mix adjust outcomes. However, the current HOS case-mix methodology detects a greater number of positive plan outliers, which is of great value for quality improvement activities, plan accountability and public reporting. Further studies should examine what differences in care structures and processes contribute to better or worse outcomes in those high-performing plans for purposes of ‘best practices.’

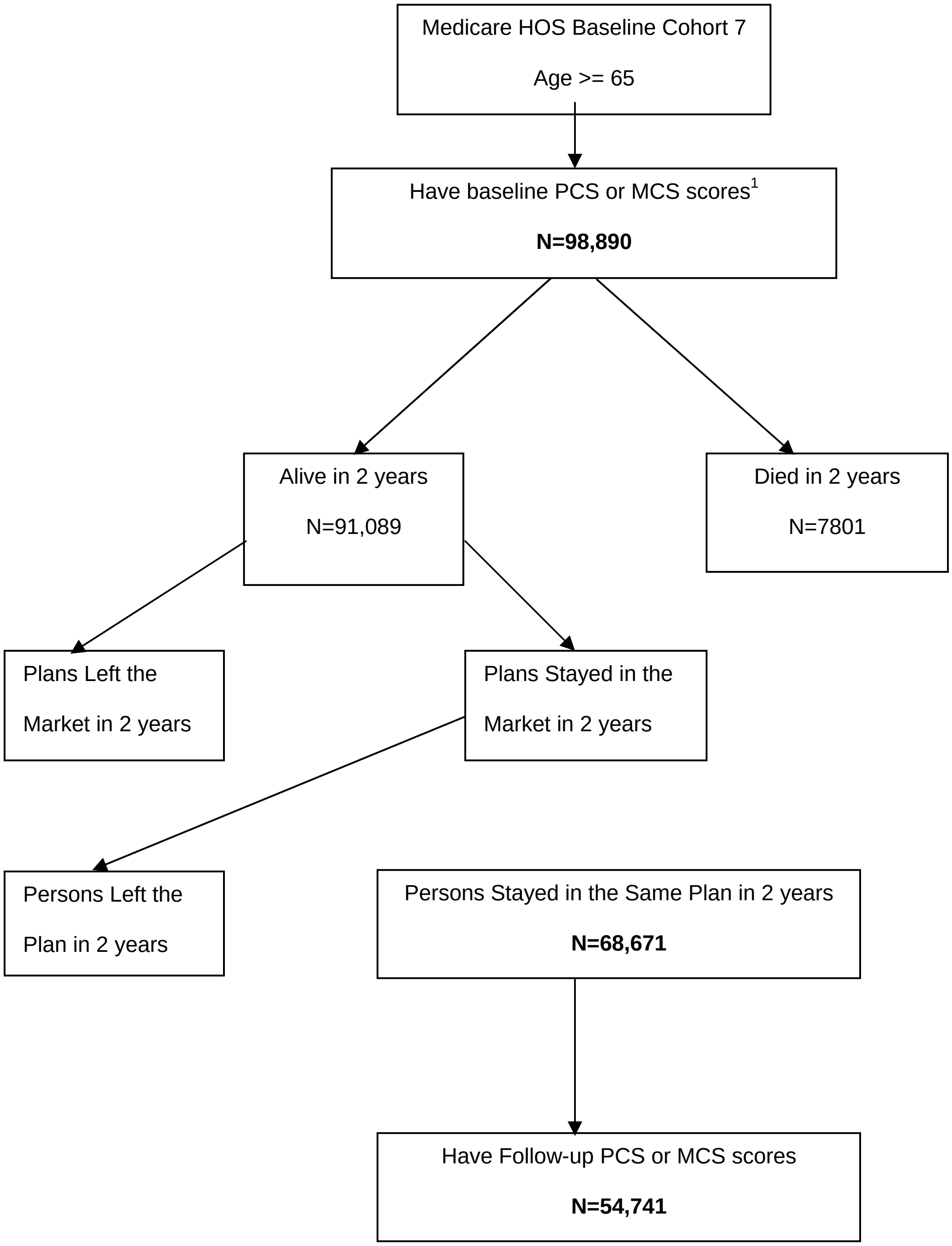
6.0 RECOMMENDATIONS FOR CMS

On the basis of the findings of this report we recommend the continued adoption of the current HOS case-mix methodology because of its conceptual relevance and ability to detect more completely high-performing plans that appears lacking in the alternative case-mix methodology that we evaluated.

A related question is whether the primary focus of measuring and reporting case-mix adjusted health status outcomes should be for public accountability of those plans serving Medicare beneficiaries, and whether the goals of this effort should be expanded to include continuous quality improvement (CQI). In our study, the analytic methods we used focused on identifying "outliers," or plans performing significantly above or below standard or average care. Thus these methods emphasize public accountability of plans, although they could also have the effect of promoting quality improvement among the plans identified as negative outliers. It can also identify those plans that are high-performers for purposes of demonstrating 'best practices' to be emulated by other plans performing less well. Public release of quality performance data has been shown in some cases to motivate hospitals, physician groups, and individual providers to take concrete steps to improve their quality of care.[31] The recommendation for continued use of the current HOS case-mix methodology is based upon the focus of the goals of the CMS HOS program for public accountability through the use of these outcome measures. Broadening of these goals for purposes of CQI should also be strongly considered.

7.0 TABLES AND FIGURES

Figure 1 Medicare Health Outcome Study (HOS) Cohort 7



[1]They had enough data to score the physical (PCS) or mental (MCS) component summary scores using missing data regression estimation (MRE) algorithms.

Table 1 List of the Case-Mix Variables by Domains for the New Case-Mix Methodology and the Currently Used HOS Case-Mix Methodology

	New Case-Mix Methodology	**HOS Case-Mix Methodology**	
Domains	**Case-Mix Variables**	**Case-Mix Variables for Mortality**	**Case-Mix Variables for PCS and MCS**
Socio-demographic	Age (linear) Gender Race/Ethnicity (Whites, African-American, Hispanic and others) Married (or not married) High school graduate (or not high school graduate) Income less than $20,000 (or income of $20,000 or greater) On Medicaid (or not on Medicaid)	Age (linear), Age 75+, Age 85+ (Models A-H) Gender (Models A-H) Race/Ethnicity (Black/African-American, Hispanic, Asian/Pacific Islander) – (Models A-H) Married or single, divorced, widowed, separated (Models A-D) High school graduate or not high school graduate (Models A-D) Income less than $20,000 or reported income of $20,000 or greater (Models A, C) On Medicaid or not on Medicaid (Models A-H) Age and Gender interaction (Models A-H) Home owner or non-home owner (Models A-D)	Age (linear), Age 75+, Age 85+ (Models A-C) Gender (Models A-C) Race/Ethnicity (Black/African-American, Hispanic, Asian/Pacific Islander) – (Models A-C) High school graduate or not high school graduate (Models A-B) Married or single, divorced, widowed, separated (Models A-B) Income less than $20,000 or reported income of $20,000 or greater (Model A) On Medicaid or not on Medicaid (Models A-C) Age and Gender interaction (Models A-C) Home owner or non-home owner (Models A-B)

Table 1 List of the Case-Mix Variables by Domains for the New Case-Mix Methodology and the Currently Used HOS Case-Mix Methodology

	New Case-Mix Methodology	**HOS Case-Mix Methodology**	
Domains	**Case-Mix Variables**	**Case-Mix Variables for Mortality**	**Case-Mix Variables for PCS and MCS**
Co - morbidities	Presence or absence of each of 13 chronic conditions: hypertension, acute myocardial infarction/angina, coronary artery disease, congestive heart failure, other heart conditions, stroke, pulmonary disease, gastrointestinal disorders, arthritis of hip or knee, arthritis of hand or wrist, sciatica, diabetes, cancer other than skin cancer. Treatment or non-treatment for 4 cancer types: colon/rectal, lung, breast, prostate	Presence or absence of each of 13 chronic conditions: hypertension, acute myocardial infarction/angina, coronary artery disease, congestive heart failure, other heart conditions, stroke, pulmonary disease, gastrointestinal disorders, arthritis of hip or knee, arthritis of hand or wrist, sciatica, diabetes, cancer other than skin cancer (Models A, B) Treatment or non-treatment for 4 cancer types: colon/rectal, lung, breast, prostate (Models A,B)	
Baseline Health Status	Baseline PCS and MCS scores (linear)	SF-36 Physical Functioning/Activities of Daily Living Index (Models A-E) SF-36 General Health scale (Models A-E) SF-36 Social Functioning scale (Models A-E) One-item measure of General Health compared to others (Models A-E) Baseline PCS and MCS scores (linear) (Models F-G)	

Table 2 Patient Characteristics in the Medicare Advantage Program

	Medicare Advantage Program (N=147,955)
Age, years (Mean±SD)	76.3 (±6)
65 – 74	46.3%
75 – 84	40.1%
85+	13.6%
Gender –male	40.8%
Race – Whites	85.1%
- African Americans	9.3%
- Hispanics	2.5%
- Others	3.1%
Married	54.9%
Education <12 years	29.7%
Income <$20,000	47.3%
Home Owner	72.1%
Medicaid Eligible	7.4%
Number of Comorbidities (Mean ± SD)	2.8 (+2)
Comorbidity 0	10.3%
1	17.9%
2	20.2%
3	18.2%
>=4	33.3%
Hypertension	62.0%
Diabetes	20.1%
Angina/coronary artery disease	16.2%
Acute myocardial infarction	11.2%
Congestive heart failure	9.0%
Other heart conditions	22.6%
Stroke	9.2%
Pulmonary disease	13.7%
Gastrointestinal disorders	5.0%
Arthritis of hip or knee	44.9%
Arthritis of hand or wrist	38.5%
Sciatica	23.0%
Cancer other than skin cancer	14.9%
Treatment for colon/rectal cancer	1.1%
Treatment for lung cancer	0.6%
Treatment for breast cancer	1.9%
Treatment for prostate cancer	3.0%
Baseline Physical Health (Mean ± SD)*	39.6 (+12)
Baseline Mental Health (Mean ± SD)*	51.9 (+10)

*Scale from 0 (=worst health) to 100 (=best health).

Table 3 The C-Statistic and Hosmer–Lemeshow Results for Sequential Models

Predicting Mortality, PCS the same or better and MCS the same or better in the derivation sample.

Models	Mortality			PCS Same or Better			MCS Same or Better		
	C-statistic	Hosmer–Lemeshow	Pseudo R-square	C-statistic	Hosmer–Lemeshow	Pseudo R-square	C-statistic	Hosmer–Lemeshow	Pseudo R-square
Sociodemographics	0.71	P = 0.033	0.080	0.53	P = 0.018	0.002	0.57	P = 0.087	0.009
Sociodemographics + Medical Conditions	0.77	P = 0.018	0.138	0.54	P = 0.066	0.003	0.58	P = 0.007	0.013
Sociodemographic + Medical Conditions + Baseline PCS&MCS	0.80	P = 0.903	0.178	0.67	P = 0.214	0.0623	0.69	P = 0.415	0.076

Covariates used in the models: Sociodemographics (age (linear), gender, race/ethnicity, marital status, level of education (<12 years), and income (<$20 000)), presence or absence of each of 13 chronic conditions: hypertension, acute myocardial infarction/angina, coronary artery disease, congestive heart failure, other heart conditions, stroke, pulmonary disease, gastrointestinal disorders, arthritis of hip or knee, arthritis of hand or wrist, sciatica, diabetes, cancer other than skin cancer, treatment or non-treatment for 4 cancer types: colon/rectal, lung, breast, prostate, baseline PCS and MCS scores.

C-statistic reflects the predictive power of the models to discriminate among patients by ordering them according to rates of the outcome event. A c-statistic value of less than 0.5 indicates poor discriminatory power of the model.

Table 4 Case-Mix Models for Mortality, PCS the Same or Better and MCS the Same or Better

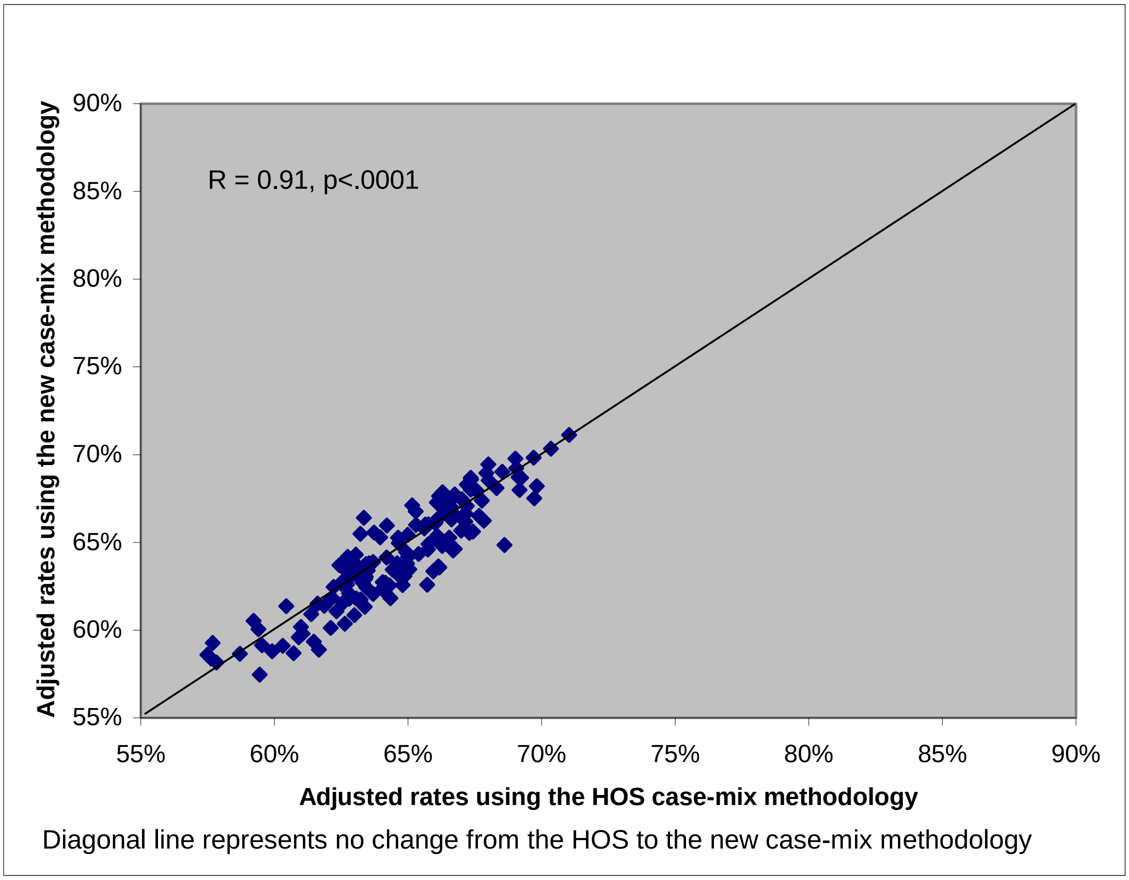
	Mortality								
	B coeff	Odd Ratio (95% CI)		B coeff	Odd Ratio (95% CI)		B coeff	Odd Ratio (95% CI)	
Age	0.0865	1.090	1.084 - 1.097	-0.0361	0.965	0.960 - 0.969	-0.0178	0.982	0.977 - 0.988
Female	-0.5058	0.603	0.550 - 0.662	-0.0336	0.967	0.908 - 1.030	0.0120	1.012	0.943 - 1.087
Married	-0.2079	0.812	0.739 - 0.893	-0.1044	0.901	0.844 - 0.962	-0.1390	0.870	0.808 - 0.937
African-Americans	0.0455	1.047	0.890 - 1.230	0.1084	1.115	0.984 - 1.263	-0.0956	0.909	0.796 - 1.038
Asian	-0.3298	0.719	0.487 - 1.061	0.1031	1.109	0.859 - 1.430	-0.3648	0.694	0.539 - 0.895
Hispanics	-0.2592	0.772	0.563 - 1.059	-0.0474	0.954	0.742 - 1.226	-0.2019	0.817	0.629 - 1.062
Education <12	0.0219	1.022	0.934 - 1.118	-0.0429	0.958	0.894 - 1.026	-0.2581	0.772	0.717 - 0.832
Income < $20,000	-0.0393	0.961	0.877 - 1.053	-0.1310	0.877	0.821 - 0.937	-0.1650	0.848	0.789 - 0.912
Medicaid eligibility	0.1243	1.132	0.970 - 1.321	-0.1152	0.891	0.773 - 1.028	-0.2374	0.789	0.682 - 0.912
Hypertension	-0.1862	0.830	0.761 - 0.906	-0.1536	0.858	0.808 - 0.911	0.0709	1.073	1.004 - 1.148
Diabetes	0.2309	1.260	1.143 - 1.388	-0.2195	0.803	0.745 - 0.866	-0.0798	0.923	0.851 - 1.002
Acute Myocardial Infarction	0.0559	1.058	0.932 - 1.200	-0.2458	0.782	0.700 - 0.874	0.0876	1.092	0.966 - 1.233
Coronary Artery Disease	-0.0429	0.958	0.853 - 1.076	-0.0976	0.907	0.826 - 0.996	0.0312	1.032	0.932 - 1.142
Congestive Heart Failure	0.6293	1.876	1.664 - 2.116	-0.1717	0.842	0.743 - 0.955	0.0657	1.068	0.934 - 1.221
Other heart conditions	0.0254	1.026	0.931 - 1.130	-0.1090	0.897	0.833 - 0.966	-0.0162	0.984	0.908 - 1.067
Stroke	0.4721	1.603	1.436 - 1.790	-0.2560	0.774	0.694 - 0.864	-0.1339	0.875	0.779 - 0.983
COPD – Asthma	0.3663	1.442	1.302 - 1.598	-0.3285	0.720	0.660 - 0.786	0.00975	1.010	0.919 - 1.109
Gastro-intestinal Disorder	-0.1172	0.889	0.747 - 1.059	-0.0767	0.926	0.809 - 1.060	-0.0300	0.970	0.837 - 1.124
Arthritis Hip	-0.2509	0.778	0.708 - 0.855	-0.3557	0.701	0.656 - 0.749	0.1113	1.118	1.039 - 1.203
Arthritis Hand	-0.2974	0.743	0.676 - 0.817	-0.1057	0.900	0.842 - 0.961	-0.00232	0.998	0.928 - 1.073

Table 4 Case-Mix Models for Mortality, PCS the Same or Better and MCS the Same or Better

	Mortality			PCS the Same or Better			MCS the Same or Better		
	B coeff	**Odd Ratio**	**(95% CI)**	**B coeff**	**Odd Ratio**	**(95% CI)**	**B coeff**	**Odd Ratio**	**(95% CI**
Sciatica	-0.3138	0.731	0.661 - 0.808	-0.0803	0.923	0.858 - 0.992	-0.0245	0.976	0.903 - 1.055
Cancer	0.5559	1.744	1.548 - 1.963	-0.0334	0.967	0.874 - 1.070	0.0748	1.078	0.963 - 1.206
Treatment for colon cancer	0.0587	1.060	0.756 - 1.487	-0.1311	0.877	0.614 - 1.252	-0.2246	0.799	0.551 - 1.158
Treatment for prostate cancer	-0.3906	0.677	0.472 - 0.970	-0.2302	0.794	0.602 - 1.048	-0.0287	0.972	0.715 - 1.321
Treatment for lung cancer	-0.2284	0.796	0.633 - 1.001	-0.1000	0.905	0.734 - 1.116	-0.0944	0.910	0.722 - 1.148
Treatment for breast cancer	1.7596	5.810	4.108 - 8.217	-0.1021	0.903	0.489 - 1.668	-0.0370	0.964	0.504 - 1.844
Baseline Physical Health (PCS)	-0.0463	0.955	0.951 - 0.959	-0.0672	0.935	0.932 - 0.938	0.0457	1.047	1.043 - 1.050
Baseline Mental Health (MCS)	-0.0301	0.970	0.967 - 0.974	0.0182	1.018	1.015 - 1.022	-0.0587	0.943	0.939 - 0.947

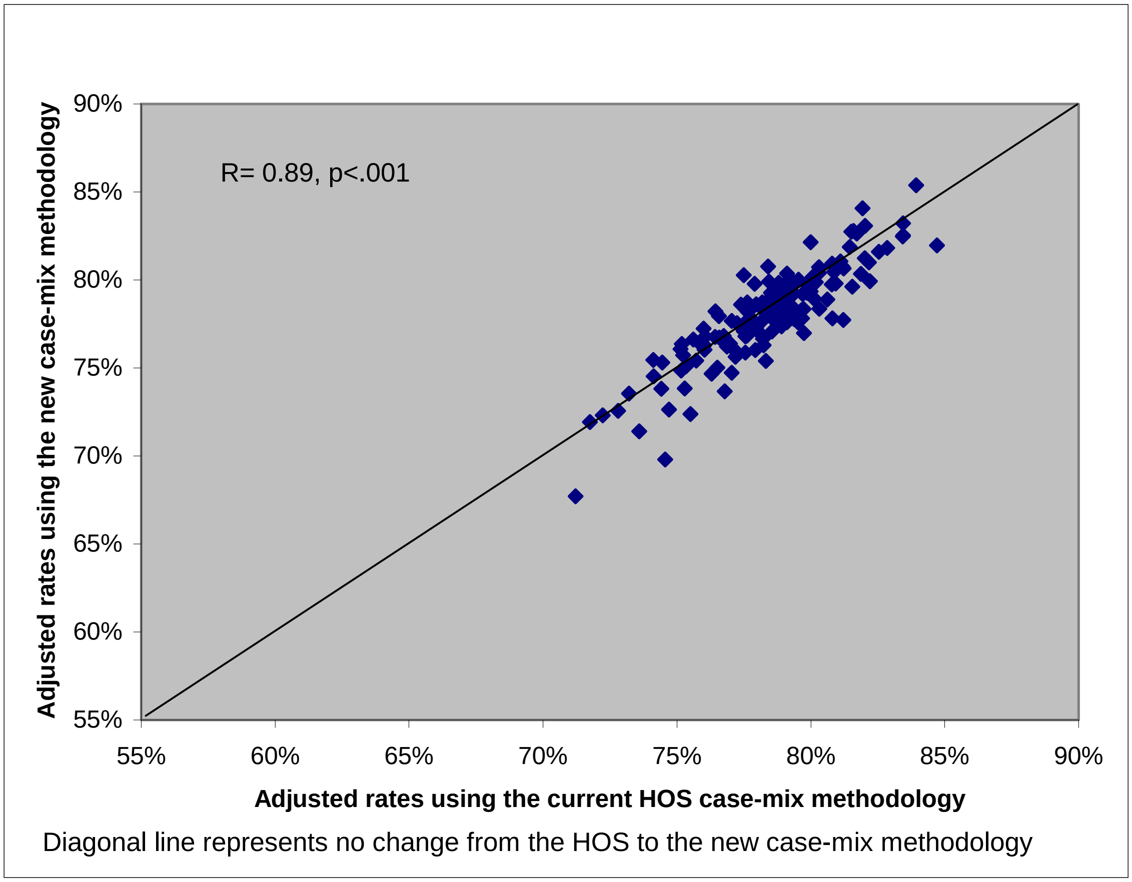
Reference groups = male, not married, white, Education $\geq$ 12, Income $\geq$ \$20,000

Figure 2 Comparing the Plan Rates Between the New Case-Mix Methodology and the HOS Case-Mix Methodology for Being Alive and PCS the Same or Better



Number of plans = 150

Figure 3 Comparing the Plan Rates Between the New Case-Mix Methodology and the HOS Case Mix Methodology for MCS the Same or Better



Number of plans = 150

Table 5 Comparing the Plan Outliers Between the New Case-Mix Methodology and the HOS Case-Mix Methodology

Outcome: Alive with PCS the same or better			
Positive outliers		**Negative outliers**	
New Case-Mix Methodology	HOS Case-Mix Methodology	New Case-Mix Methodology	HOS Case-Mix Methodology
H3653	H3653 H2261 H3204 H3864 H3907 H4003 H4506	H0630 H2931 H3749 H5005 H0545 H1349 H1555 H2462	H0630 H2931 H3749 H5005 H2949 H3164 H3312

Outcome: MCS the same or better			
Positive outliers		**Negative outliers**	
New Case-Mix Methodology	HOS Case-Mix Methodology	New Case-Mix Methodology	HOS Case-Mix Methodology
H0657 H1463 H3362 H5211	H0657 H1463 H3362 H5211 H0502 H0752 H2459 H3954	H1019 H1099 H3204 H4506 H4003	H1019 H1099 H3204 H4506

Number of plans = 150

Appendix A Description of the Current HOS Case-Mix Models

DEATH MODEL	DESCRIPTION	N*	Pseudo R-square	C-statistic	H-L test
A	Socio-demographic variables (including poverty status), baseline health status measures, and individual chronic conditions.	67,756	0.191	0.810	0.5244
B	Same model as Model A, without the poverty variable (information on income was missing in many surveys).	82,549	0.193	0.812	0.5106
C	Same as Model A, but uses disease groupings (e.g., Mci) rather than individual disease coefficients.	76,501	0.184	0.806	0.8412
D	Same as Model C, without the poverty variable.	93,833	0.187	0.808	0.3751
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.	100,187	0.186	0.807	0.0645
F	Same as Model E, but substitutes baseline PCS and MCS for the other functioning variables (PFADL, C7GH, C7SF, C7CMPHTH).	102,856	0.167	0.797	0.3752
G	CMS demographic variables and baseline PCS and MCS only.	104,959	0.140	0.773	0.4468
H	CMS demographics only.	106,057	0.078	0.705	0.0165

Appendix A Description of the Current HOS Case-Mix Models

PCS SAME BETTER MODEL	DESCRIPTION	N*	Pseudo R-square	C-statistic	H-L test
A	Sociodemographic variables (including poverty Status)	44,828	0.002	0.532	0.2875
B	Same model as Model A, without the poverty variable.	54,108	0.002	0.533	0.4394
C	Only CMS variables (age, gender and Medicaid status).	58,065	0.002	0.533	0.2799

MCS SAME BETTER MODEL	DESCRIPTION	N*	Pseudo R-square	C-statistic	H-L test
A	Sociodemographic variables (including poverty status)	44,736	0.008	0.565	0.0006
B	Same model as Model A, without the poverty variable.	53,976	0.007	0.562	<.0001
C	Only CMS variables (age, gender and Medicaid status).	57,912	0.005	0.548	<.0001

*Medicare-HOS Cohort 7

Appendix B Comparing the Plan Outliers Among the New Case-Mix Adjusted Models With and Without Baseline PCS and MCS Scores and the HOS Case-Mix Methodology

Outcome: Alive with PCS the Same or Better					
Positive Outliers			**Negative Outliers**		
New Case-Mix Methodology without PCS and MCS Baseline Scores	New Case-Mix Methodology with PCS and MCS Baseline Scores	HOS Case-Mix Methodology	New Case-Mix Methodology without PCS and MCS Baseline Scores	New Case-Mix Methodology with PCS and MCS Baseline Scores	HOS Case-Mix Methodology
H3653	H3653	H3653 H2261 H3204 H3864 H3907 H4003 H4506	H0630 H2931 H3749 H5005 H2462 H3312	H0630 H2931 H3749 H5005 H2462 --- H0545 H1349 H1555	H0630 H2931 H3749 H5005 --- H3312 H2949 H3164

Appendix B Comparing the Plan Outliers Among the New Case-Mix Adjusted Models With and Without Baseline PCS and MCS Scores and the HOS Case-Mix Methodology

Outcome: MCS the Same or Better					
Positive Outliers			**Negative Outliers**		
New Case-Mix Methodology without PCS and MCS Baseline Scores	New Case-Mix Methodology with PCS and MCS Baseline Scores	HOS Case-Mix Methodology	New Case-Mix Methodology without PCS and MCS Baseline Scores	New Case-Mix Methodology with PCS and MCS Baseline Scores	HOS Case-Mix Methodology
--- --- H3362 --- H0752	H0657 H1463 H3362 H5211	H0657 H1463 H3362 H5211 H0752 H0502 H2459	H1019 H1099 H3204 H4506	H1019 H1099 H3204 H4506 H4003	H1019 H1099 H3204 H4506

Number of plans = 150

8.0 REFERENCES

[1] HEDIS®2003 Volume 6. February 2003. Specifications for the Medicare Health Outcomes Survey Washington, D.C.: National Committee for Quality Assurance.

[2] Brook RH, McGlynn EA, Cleary PD. Quality of health care. Part 2: measuring quality of care. N Engl J Med. 1996;335:966-70.

[3] Jencks SF, Dobson A. Refining case-mix adjustment. The research evidence. N Engl J Med. 1987;317:679-86.

[4] Evaluating the Two-Year Follow-up Health Status of Medicare Fee-For –Service Beneficiaries Using the Health Outcomes Survey. http://www.hosonline.org/surveys/hos/download/HOS_FFS_2YR_Health_Status.pdf. Accessed July 18, 2007.

[5] Calculating Medicare Health Outcomes Survey Performance Measurement Results. http://www.hosonline.org/surveys/hos/download/HOS_Calculating_PM_Results.pdf. Accessed July 18, 2007.

[6] Kazis LE, Ren XS, Lee A, et al. Health status in VA patients: Results from the Veterans Health Study. Am J Med Qual. 1999;14:28-38.

[7] Kazis LE, Selim A, Rogers W, Ren XS, Lee A, Miller DR. Dissemination of methods and results from the veterans health study: final comments and implications for future monitoring strategies within and outside the veterans healthcare system. J Ambul Care Manage. 2006;29:310-9.

[8] The Veterans RAND 12 Item Health Survey: What it is and how it is used. http://www.va.gov/chqoer/VR12.pdf. Accessed July 30, 2007.

[9] Imputing the Physical and Mental Summary Scores (PCS and MCS) for the MOS SF-36 and the Veterans SF-36 Health Survey in the presence of Missing Data. http://www.hosonline.org/surveys/hos/download/HOS_Veterans_36_Imputation.pdf. Accessed July 30, 2007.

[10] Imputing Physical and Mental Summary Scores (PCS and MCS) for the Veterans SF-12 Health Survey in the Context of Missing Data.

http://www.hosonline.org/surveys/hos/download/HOS_Veterans_12_Imputation.pdf. Accessed July 18, 2007.

[11] Kazis LE, Miller DR, Skinner KM, et al. Patient-reported measures of health: The Veterans Health Study. J Ambul Care Manage. 2004;27:70-83.

[12] Ware JE, Kosinski M, Keller SK. SF-36 Physical and Mental Health Summary Scales: A Users's Manual Boston, MA, The Health Institute, New England Medical Center 1994.

[13] Ware JE, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scores. The Health Institute, New England Medical Center, Boston, Massachusetts, March, 1995.

[14] Jones D, Kazis L, Lee A, Rogers W, Skinner K, Cassar L, Wilson N, Hendricks A. Health status assessments using the Veterans SF-36 and SF-12. Methods for evaluating outcomes in the Veterans Health Administration. Journal of Ambulatory Care Management 2001; 24(3):1-19.

[15] Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. JAMA. 1995;273:59-65.

[16] Hornbrook MC, Goodman MJ. Chronic disease, functional health status, and demographics: a multi-dimensional approach to risk adjustment. Health Serv Res. 1996;31:283-307.

[17] Williams DR. Race/ethnicity and socioeconomic status; measurement and methodological issues. Int J Health Serv. 1996;26:483-505.

[18] Bayliss EA, Bayliss MS, Ware JE Jr, Steiner JF. Predicting declines in physical function in persons with multiple chronic medical conditions: what we can learn from the medical problem list. Health Qual Life Outcomes. 2004;2:47.

[19] Dorr DA, Jones SS, Burns L, Donnelly SM, Brunker CP, Wilcox A, Clayton PD. Use of health-related, quality-of-life metrics to predict mortality and hospitalizations in community-dwelling seniors. J Am Geriatr Soc. 2006;54:667-73.

[20] Kazis LE, Miller DR, Skinner KM, Lee A, Ren XS, Clark JA, Rogers WH, Spiro A III, Selim AJ, Linzer M, Payne S, Mansell D, Fincke G. Patient reported measures of health: The Veterans Health Study. Journal of Ambulatory Care Management. 2004; 27:1, 70-83.

[21] Joint Commission on Accreditation of Healthcare Organizations. Ambulatory Health Care Standards Manual. Chicago, IL: JCAHO; 1990.

[22] Bayliss EA, Bayliss MS, Ware JE, Steiner JF. Predicting declines in physical function in

persons with multiple chronic medical conditions: what we can learn from the medical problem list? Health Qual Life Outcomes 2004; 2: 47.

[23] Stewart AL, Greenfield S, Hays RD et al. Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. JAMA 1989; 262: 907–913.

[24] Thomas JW, Holloway JJ, Guire KE. Validating risk-adjusted mortality as an indicator for quality of care. Inquiry. 1993; 30: 6–22.

[25] Best WR, Cowper DC. The ratio of observed-to-expected mortality as a quality of care indicator in non-surgical VA patients. Med Care. 1994; 32: 390–400.

[26] Tierney WM, Takesue BY, Vargo DL, et al. Using electronic medical records to predict mortality in primary care patients with heart disease: prognostic power and pathophysiologic implications. J Gen Intern Med. 1996; 11: 83–91.

[27] Selim AJ, Berlowitz D, Fincke G, Rogers W, Qian S, Lee A, Cong Z, Selim BJ, Ren XS, Rosen AK, Kazis LE. Use of risk-adjusted change in health status to assess the performance of integrated service networks in the Veterans Health Administration. Int J Qual Health Care. 2006;18:43-50.

[28] Oldridge N, Gottlieb M, Guyatt G et al. Predictors of health related quality of life with cardiac rehabilitation after acute myocardial infarction. J Cardiopulm Rehabil. 1998;18:95–103.

[29] Ware JE, Bayliss MS, Rogers WH et al. Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems. JAMA 1996; 276:1039–1047.

[30] Stewart AL, Greenfield S, Hays RD, Wells K, Rogers WH, Berry SD, McGlynn EA, Ware JE. Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. JAMA. 1989;262:907-913.

[31] Shahian D, Normand S, Torchiana D, Lewis S, Pastore J, Kuntz R, Dreyer. P. Cardiac surgery report cards: comprehensive review and statistical critique. Annals of Thoracic Surgery. 2001; 72:2155-2168.