NCQA

Measuring the Quality of America's Health Care

IMPLEMENTING THE HEDIS®
MEDICARE HEALTH OUTCOMES SURVEY

The Impact of Health Plan Quality on
Medicare Beneficiary Outcomes

TECHNICAL REPORT
MAY 21, 2007

Judy H. Ng, PhD
Sarah Hudson Scholle, DrPH
L. Gregory Pawlson, MD, MPH
The National Committee for Quality Assurance (NCQA)
Washington, DC

Jeffrey S. Harman, PhD
The Department of Health Services Research, Management and Policy
College of Public Health and Health Professions
University of Florida, Gainesville

Russell E. Mardon, PhD
Westat
Rockville, MD

Arlene S. Bierman, MD, MS
Centre for Research in Inner City Health
University of Toronto, Canada

Questions concerning this work may be emailed to Dr. Ng at: ng@ncqa.org

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

The Impact of Health Plan Quality on
Medicare Beneficiary Outcomes

Table of Contents

1.0   EXECUTIVE SUMMARY

This report assesses the relationship of plan-level performance on Health Employer Data and Information Set (HEDIS®) measures of clinical processes and intermediate outcomes with changes over two-years in the self-reported physical and mental health outcomes from the Medicare Health Outcomes Survey (HOS) among elderly Medicare plan enrollees with diabetes, hypertension, ischemic heart disease, and depression.

Data are from the plan-level HEDIS measures for 2002 related to diabetes, hypertension, ischemic heart disease and depression care and patient-level longitudinal 2001-2003 health outcomes as measured by HOS physical and mental scores captured by the Medical Outcomes Study Short Form-36 (MOS SF-36®). To account for clustering of individuals within health plans, hierarchical linear models estimated the imputed effects of plan performance of HEDIS measures on individual enrollees' longitudinal health outcomes measured by the MOS SF-36. The unit of analysis was the individual.

Results indicate that plan performance on HEDIS measures was related to significant positive improvements in health outcomes of enrollees with diabetes. For enrollees with diabetes, plan performance on HEDIS intermediate outcome measurement was significantly and positively related to improvements in enrollee physical health functioning Plan performance on process measures in diabetes was associated with improvements in enrollee mental health functioning. We did not find significant associations for the other conditions of hypertension, ischemic heart disease, and depression studied. We believe that this failure to show a significant effect may be related to the relatively short time frames of the current study, and issues related to the quality of the process measurements and differences in the definitions of persons with depression and heart disease included in the HOS sample.

This study represents one of the first attempts to directly link plan HEDIS performance to outcome measures of enrollee health. Both HEDIS process and intermediate outcomes measures for diabetes care are positively related to enrollee health outcomes, suggesting a possible direct

linkage of HEDIS measures of process and intermediate outcomes with the outcome measures of mental and physical functioning included in the HOS survey.

2.0   INTRODUCTION

The National Committee for Quality Assurance (NCQA) began accrediting managed care organizations in 1991. The purpose of this accreditation process is to allow consumers, employers, regulators, and health plans to view objective data on the quality of managed care organizations. In addition to accreditation, NCQA developed an ever enlarging set of standardized clinical performance measures, which were entitled the "Health Plan Employer Data and Information Set," or as it is widely known, HEDIS®.[1] These measures have been widely adopted by plans, are required by many private, state and federal entities and in short, are seen by many as the gold standard for clinical quality measurement. In 1999, NCQA began requiring health plans that were undergoing accreditation by NCQA to report HEDIS measures, and also became the first (and to date the only) accrediting body that actually incorporates the relative performance of health plans on clinical measures as a core component of their accreditation score.

Despite substantial progress in measuring and reporting quality in health care using indicators such as HEDIS, there are relatively limited data directly linking any clinical process measures of the use of effective clinical services (e.g., screening for hemoglobin A1c in diabetics) or intermediate outcomes (e.g., control of hemoglobin bA1c in diabetics) with endpoints of care such as functional outcomes, health status or mortality. While these health outcomes are critical indicators of the end results of healthcare,[2] they are often difficult to link directly to the actions of healthcare providers and are not in and of themselves, actionable by providers.[3] [4] Thus most quality evaluations use measures of structure, process or intermediate outcomes, as for example, HbA1c screening and control in diabetics, cholesterol screening and control post myocardial infarction (MI). [5] [6]   However, for full validity as measures of quality, measures of structure, process or intermediate outcomes should be positively linked to global health outcomes such as health status, which are of direct relevance and are more meaningful from the patient perspective.[7] [8] [9]

Even more importantly, there is a special need for research that better ties clinical processes to patient outcomes in the Medicare population, where given the limited reserve of older patients,

management of chronic conditions is likely to have a major effect on quality of life. While global health outcomes such as mortality and functional status are considered to be the most critical indicators of healthcare quality,[10] as noted above, these types of outcomes are often difficult to link to the actions of healthcare providers.[11,12] Based on published literature, the only joint analysis of Medicare Health Outcomes Survey (HOS) data and HEDIS clinical data found a positive correlation between several clinical measures and enrollee health status.[13] However, the study did not consider longitudinal outcomes and did not control for possible differences in health plan populations.

This study looked for a linkage between clinical processes and outcomes on a more longitudinal basis correcting for health plan population differences, using data from two unique databases overseen by NCQA, HEDIS and the Medicare Health Outcomes Survey (HOS).[14] HEDIS clinical performance measurement and HOS measurement of patient-specific functional status change over time are required parts of health plan quality reporting in the Medicare Advantage (MA) program. Thus, MA offers an opportunity to assess correlations of health plans' clinical performance using process and intermediate outcome measures, and individual enrollee's health status.

We merged plan level data on quality with member level outcomes data to assess the association of changes in health outcomes among individual Medicare beneficiaries with their health plan's performance on relevant HEDIS measures for diabetes, hypertension, ischemic heart disease, and depression. This is the first study to merge these two large-scale national databases and analyze them jointly for disease-specific subgroups. We chose enrollees with diabetes, hypertension, ischemic heart disease and depression not only because these are highly prevalent, costly conditions requiring continuous maintenance treatment, but there are numerous HEDIS indicators relating to the conditions' care and outcomes.[15] By restricting each model to beneficiaries with a specific common chronic condition, the analysis can better pinpoint the link between measures of clinical processes (for example, screening for cholesterol in patients post-myocardial infarction) and intermediate outcomes (such as control of cholesterol levels) and functional outcomes. We hypothesized that enrollees in health plans with high performance on

relevant HEDIS measures would have better patient-reported health outcomes than enrollees in lower performing plans.

3.0   METHODS

3.1   Data

This study used data from the HEDIS[16] and the HOS version 1. [17] HEDIS is the most widely used set of clinical performance measures in the managed care industry. The HOS version 1 is a health survey instrument developed to analyze the impact of Medicare health plans on self-reported individual health status, as measured by the Medical Outcomes Study Short Form-36 (MOS SF-36®).[18] The usefulness of the MOS SF-36 in measuring changes in outcomes has been well-documented for a variety of conditions,[19,20,21,22] including diabetes,[23][24][25][26] hypertension,[27][28] ischemic heart disease,[29][30][31] and depression.[32][33][34][35] The HOS protocol includes mailings with telephone follow-up.[36]

This study uses Cohort 4 of HOS, which collected baseline in 2001 and follow-up survey data in 2003, the most recent years of baseline and follow-up data available at the time of analysis. All non-institutionalized, non-proxy respondents age 65 or greater, who returned a usable baseline survey and who either returned a usable follow-up survey, or who died before they could be resurveyed were included in the study. For this study we used HEDIS clinical data covering services provided in 2002, the year between the baseline and follow-up HOS data. The baseline 2001 HOS survey response rate was 68.4%; the follow-up 2003 survey response rate was 78.6%. The total number of health plans is 152.

3.2   Study Populations

This study focused on Cohort 4 HOS respondents with the following four self-identified conditions: diabetes (N=8,184), hypertension (N=27,206), ischemic heart disease (N=9,125), and depression (N=4,911). The samples for each disease-specific sub-population were identified based on responses to relevant items on the HOS questionnaire: 1) for diabetes, enrollees who responded 'yes' to a question asking if a doctor ever told them that they had diabetes, high blood sugar, or sugar in the urine; 2) for hypertension, enrollees who responded 'yes' to a question asking if a doctor ever told them that they had hypertension or high blood pressure; 3) for ischemic heart disease, enrollees who responded 'yes' to a question asking if a doctor ever told them that they had a myocardial infarction, heart attack, angina pectoris or coronary artery

disease; and 4) for depression, enrollees who responded that they felt depressed or sad much of the time. Table 1 illustrates the number of HOS Cohort 4 individual responses for each disease subgroup, and their percentages relative to the total number of eligible responses (all non-institutionalized, non-proxy respondents age 65 or greater, who returned a usable baseline survey and who either returned a usable follow-up survey, or who died before they could be resurveyed, n=48,521). Over half of HOS respondents reported being told they had hypertension, compared to 17% with diabetes, 19% with ischemic heart disease, and 10% with depression.

3.3   Dependent Variables

The dependent variables used in the analyses are change in physical and mental health during the study period of 2001-2003, based on the Physical Component Summary (PCS) and the Mental Component Summary (MCS) scores of the MOS SF-36. The PCS and the MCS are both standardized to a 0-100 scale, with 0 representing worst health and 100 representing best health. Both scores are standardized using norms for the general United States population so that a score of 50 represents the national average and the standard deviation of each score's distribution is 10.[37, 38, 39]

In order to account for enrollees who died during the study period, and to include death as an outcome, methods of adjusting the PCS and MCS for death, as described by Diehr and colleagues, were used.[40, 41, 42]   Diehr et al's methods use logistic regression to predict the probability of being healthy at follow-up (whether someone could be expected to be in the top 75th percentile on the PCS or MCS distribution at follow-up). Baseline PCS and MCS are used as the covariates to model the probability of being healthy at follow-up. Predicted values for being healthy at follow-up were then generated from the logistic regression model. For those who died within two years of baseline measurement, a value of "0" was given, indicating 0 probability of being healthy at follow up. This accounts for inflation of PCS and MCS scores that occurs if deaths before follow-up are removed from analysis.

To generate the PCS and MCS change scores, the baseline predicted value was subtracted from the follow up predicted value, and these change scores were used as the dependent variables in the analyses. Separate equations were estimated for change in PCS and change in MCS.

3.4 Independent Variables

HEDIS Measures

Table 2 shows the independent variables of interest in the analyses, based on relevant HEDIS clinical measures for the four conditions: diabetes, hypertension, ischemic heart disease and depression. The relevant HEDIS measures correspond to diabetes management and control for enrollees with diabetes (six measures), hypertension control for enrollees with hypertension (one measure), follow-up after acute cardiovascular events for enrollees with ischemic heart disease (three measures), and treatment of depression for enrollees with depression (five measures). The relevant HEDIS measures include both process and intermediate outcomes measures for most of the four conditions.

To simplify the analysis for diabetes, the six HEDIS measures were grouped into two composite measures: a process composite measure and an intermediate outcomes composite measure. The process composite measure was equal to the mean performance of the health plan on the four process-related performance measures (Hba1c testing, eye exams, LDL screenings, and nephropathy monitoring). The intermediate outcomes composite measure was equal to the mean performance of the health plan on the two intermediate outcomes-related performance measures (HbA1c control and LDL control). Because the HEDIS criteria for HbA1c was originally for poor control, the percent in poor control was inverted - subtracted from one - so that this measure reflected the proportion of enrollees with positive outcomes, corresponding to the proportion of enrollees with positive outcomes as represented in other HEDIS diabetes measures.

The three HEDIS measures related to follow-up after acute cardiovascular events were: the percent of enrollees with acute myocardial infarction who receive beta-blocker prescriptions within seven days of hospital discharge; the percent of patients with cholesterol screening following an acute myocardial infarction (AMI), coronary artery bypass graft (CABG), or percutaneous transluminal coronary angioplasty (PCTA); and the percent of enrollees with LDL under control (as measured by a value of LDL of 130 or less) following an acute myocardial infarction, coronary artery bypass graft, or percutaneous transluminal coronary angioplasty.

Lastly, the five HEDIS measures related to treatment of depression were: the percent of enrollees with depression with optimal (continued) practitioner contact following diagnosis, the percent of enrollees continuing on antidepressants for 3 months following diagnosis (acute phase treatment), the percent of enrollees continuing on antidepressants for 6 months following diagnosis (continuation phase treatment), the percent of enrollees with depression who receive 7-day outpatient follow-up after hospitalization for mental illness, and the percent of enrollees with depression receiving 30-day outpatient follow-up after hospitalization for mental illness.

Individual-Level Covariates

To control for differences in case mix between plans, several individual-level covariates were included in the models. All individual covariates were measured at baseline. These individual-level covariates included demographic factors and clinical factors that have been shown to be associated with health services utilization and health outcomes. Demographic factors included enrollee age, gender, educational level (high school graduate or not), race (African American, Caucasian, or other race), ethnicity (Hispanic or not), marital status (married or not), and home ownership.

Clinical factors included the number of self-identified chronic conditions (ranging from 0 to 13) from a list used in the HOS and commonly found among Medicare beneficiaries – hypertension or high blood pressure, angina pectoris or coronary artery disease, congestive heart failure, myocardial infarction, other heart conditions, stroke, emphysema or asthma or chronic obstructive pulmonary disease, gastro-intestinal conditions, arthritis of the hip or knee, arthritis of the hand or wrist, sciatica, diabetes or high blood sugar, and cancer. Symptom severity was assessed for diabetes, hypertension and ischemic heart disease (see Table 3). Presence of depressed mood was included as a covariate for analyses of diabetes, hypertension, and ischemic heart disease subgroups, The baseline PCS and MCS scores included in the model are indicators of rank based on whether the individual scores were in the bottom third, middle third, or top third of the distribution. For example, an individual with a baseline PCS score in the bottom third of the distribution would be sicker than at least two-thirds of all other individuals included in the sample. Only baseline PCS was included in the equation examining changes in PCS and only baseline MCS was included in the equation examining changes in MCS.

3.5   Statistical Analysis

The basic analytic approach in this study was to compare changes in health status over time for health plan enrollees, as measured by physical and mental component scores of the MOS SF-36, to health plan performance on HEDIS measures related to the quality of care provided to enrollees with diabetes, hypertension, ischemic heart disease, and depression. Specifically, the association of health plans' HEDIS scores in 2002 to the change in enrollees' health status between 2001 and 2003 was assessed using hierarchical linear models that account for clustering of individuals within health plans. The unit of analysis was the individual enrollee.

This approach was selected based on the three key methodological issues which must be addressed in the statistical modeling to obtain unbiased estimates of the association of health plan HEDIS scores on changes in enrollee PCS and MCS scores. These are: (1) the simultaneity bias in the quality/outcome association; (2) health plan differences in case-mix and other salient features; and (3) the clustered sampling design in the HOS survey data.

The simultaneity bias arises because both HEDIS measure compliance and outcomes are likely to be associated with unmeasured severity of disease at the individual level. (For example, an individual with severe diabetes will be more likely to have physician visits and thus will be more likely to receive standard diabetic-related services, such as foot exams or retinal exams. Such an individual is also likely to have poorer outcomes because of disease severity.)   If the association between HEDIS performance and individual outcome were estimated using standard regression techniques, these estimates would be negatively biased. However, the primary research question is not the effect of individual HEDIS measure compliance on outcomes. Rather, it is whether plan level performance on HEDIS measures is associated with outcomes. By using the plan HEDIS rate in the equation, the confounding of individual HEDIS performance and individual outcome with unmeasured severity will no longer take place since plan-level HEDIS performance will not be correlated with any single enrollee's unmeasured (or measured) severity. Thus, unbiased coefficients for the association between HEDIS measures and outcome can be estimated.

Since the individual observations are clustered by plan, we estimated the associations using hierarchical linear models (HLMs), sometimes referred to as mixed models or multilevel models. [43] HLMs yield unbiased estimates for complex samples that occur when observations are clustered within one or more higher-level groups. The HLMs estimated control for unmeasured variables that are associated with individual outcomes and health plans, producing unbiased estimates of the association between health plan performance on HEDIS scores to individual enrollee changes in health status. Additionally, the indicator variables for MCS and PCS rank (i.e. indicators of whether the individual falls in the bottom, middle, or top third of the distribution) are included and interacted with plan HEDIS scores to determine if the impact of plan HEDIS scores on enrollee outcomes differentially effect sicker (or healthier) enrollees. Specifically, the HLMs estimated include variables for plan HEDIS performance measures (including intermediate outcomes or process composite measures for the diabetes subgroup), enrollee baseline PCS or MCS ranking, other individual-level covariates, as well as individual-level and plan-level random effects.

For example, the HLM estimated for enrollees with diabetes include variables for plan HEDIS Intermediate Outcome Composite (H-OC$_j$), plan HEDIS Process Composite(H-PC$_j$), enrollee baseline PCS or MCS ranking (PCS-rank1$_i$, PCS-rank2$_i$, MCS-rank1$_i$, MCS-rank2$_i$), other patient level covariates (X_i), plan-level fixed effects (Z_j), as well as individual-level and plan-level random effects. The top third of PCS and MCS (i.e. PCS-rank3, MCS-rank3) are excluded from the model and serve as the reference group. A separate HLM was estimated for each of the outcomes of interest (change in PCS and change in MCS). Therefore, in the example of enrollees with diabetes, four separate HLMs are estimated: (1) effect of intermediate outcomes composite on PCS; (2) effect of process composite on PCS; (3) effect of intermediate outcomes composite on MCS; and (4) effect of process composite on MCS. For example, the equation to estimate the effect of the plan HEDIS intermediate outcome composite on change in PCS will be:

$$Y_{ij} = \beta_{00} + [\beta_{10}\,H\text{-}OC_j + \beta_{20}\,H\text{-}OC_j{*}PCS\text{-}rank1_i + \beta_{30}\,H\text{-}OC_j{*}PCS\text{-}rank2_i$$
$$+ \beta_{40}PCS\text{-}rank1_i + \beta_{50}PCS\text{-}rank2_i + \beta_{60}X_i + \beta_{70}Z_i] + \mu_{0j} + e_{ij}$$

In the above equation, Y_{ij} is the change in PCS between baseline and follow-up for individual i in plan j, β_{00} is the grand mean (average change in PCS for all individuals), β_{10} represents the mean effect of the HEDIS intermediate outcomes composite on the change in PCS, β_{20} demonstrates whether the effect of the HEDIS intermediate outcomes composite is significantly different for enrollees whose baseline PCS value is in the bottom third of the distribution compared to enrollees in the top third, and β_{30} demonstrates whether the effect of the HEDIS intermediate outcomes composite is significantly different for enrollees whose baseline PCS value is in the middle third of the distribution compared to enrollees in the top third, μ_{0j} is random plan-level deviations, and e_{ij} is random person-level deviations.

Separate HLMs were also estimated for the other three disease subgroups. In the hypertension disease subgroup, two separate HLMs were estimated: (1) effect of HEDIS score (single hypertension HEDIS measure) on PCS; and (2) effect of HEDIS score on MCS. In the ischemic heart disease subgroup, six separate HLMs were estimated: (1) effect of receiving beta-blocker post hospital discharge on PCS; (2) effect of cholesterol screening following an acute cardiovascular event on PCS; (3) effect of LDL cholesterol screening following an acute cardiovascular event on PCS; (4) effect of receiving beta-blocker post hospital discharge on MCS; (5) effect of cholesterol screening following an acute cardiovascular event on MCS; and (6) effect of LDL cholesterol screening following an acute cardiovascular event on MCS. In the depression disease subgroup, ten separate HLMs were estimated: (1) effect of optimal practitioner contact on PCS; (2) effect of acute treatment on PCS; (3) effect of continuation treatment on PCS; (4) effect of 7 day follow-up on PCS; (5) effect of 30 day follow-up PCS; (6) effect of optimal practitioner contact on MCS; (7) effect of acute treatment on MCS; (8) effect of continuation treatment on MCS; (9) effect of 7 day follow-up on MCS; and (10) effect of 30 day follow-up MCS.

The models for all four conditions included individual-level covariates of age and number of chronic conditions. We used a model-building approach, and further refined the model for ischemic heart disease to include individual-level covariates of ischemic heart disease symptom severity (chest pain and shortness of breath), presence of depressed mood, and baseline PCS and MCS score. We also further refined the model for diabetes to include individual-level covariates

of diabetes symptom severity (neuropathy), presence of depressed mood, baseline PCS and MCS score, and demographic factors of gender, educational level, race, ethnicity, marital status, and homeownership.

All equations were estimated using the PROC MIXED procedure of SAS 8.2.

4.0   RESULTS

Results are presented for each of the four condition groups.

4.1   Diabetes

As shown in Table 4, the average health plan performance rate on each of the HEDIS measures ranged from a low of 57.7% for monitoring diabetic nephropathy to a high of 89.3% for screening for LDL cholesterol. Plan performance varied substantially for each of these measures, with rates on most measures ranging from a low of about 25% of enrollees to well over 90% of enrollees.

For the process composite, the average plan score was 76.4% (SD8.4%), ranging from a low of 41.3% to a high of 90.1%. The average plan score for the intermediate outcomes composite was 70.9% (SD 11.3%), with plan values ranging from a low of 14.9% to a high of 88.6%.

HEDIS Process Composite and PCS

The effect of the health plan composite scores on changes in individual outcomes as measured by PCS and MCS scores are summarized in Table 5. The health plan process composite score was associated with improvements in PCS; however the effect was not significant. The average effect of a 10 percentage point increase in the proportion of enrollees meeting the process composite was associated with a 4 percentage point increase (from the mean) in the probability of an enrollee experiencing an improvement in health as measured by the PCS. However, this effect failed to achieve statistical significance.

HEDIS Intermediate Outcomes Composite and PCS

The health plan intermediate outcomes composite score was significantly associated with changes in the PCS (Table 5). Overall, for each 10 percentage point increase in the composite intermediate outcomes measure, the probability that a person will have an improvement in health as measured by the PCS increases by 8 percentage points (p=.014). Additionally, an enrollee's baseline PCS score had a significant impact on the effect of plan performance on the intermediate outcomes composite on changes in PCS score. An individual whose baseline PCS score was in the middle third of the PCS distribution showed significantly less improvement in

PCS score (11.7 percentage point smaller improvement on average) compared to individuals in the top third of the PCS distribution (p=.012) when plans performed better on the intermediate outcomes composite. However, there was no significant difference in the impact of the intermediate outcomes composite score on change in PCS between the bottom and top third of the PCS distribution. In other words, the impact of plan performance on the intermediate outcomes composite was smallest for those enrollees who were neither the healthiest nor the sickest diabetics in the plan.

HEDIS Process Composite and MCS

The health plan process composite score for diabetes was significantly associated with changes in MCS score (Table 5). For each 10 percentage point increase in the composite process measure, the probability that a person will have an improvement in health as measured by the MCS increased by nearly 11 percentage points (p=.003). The enrollees' baseline MCS did not have a significant impact on the effect of the process composite on individual changes in MCS.

HEDIS Intermediate Outcomes Composite and MCS

When the association of health plan performance on the intermediate outcomes composite measure for diabetes on changes in individual outcomes as measured by MCS score was examined (Table 5), the intermediate outcomes composite was positively but not significantly related to changes in MCS score. Additionally, the enrollee's baseline MCS had no effect on the impact of their health plan's HEDIS performance on changes in MCS.

4.2   Hypertension

Health plan performance for hypertension control was moderately high, with an average rate of 57.5% (SD 7.1%), indicating that on average over half of patients with identified hypertension had blood pressure below the 140/90 HEDIS criterion. Health plan performance varied widely, from a low of 32.6% to a high of 71%.

Health plan performance on this measure was positively but not significantly associated with changes in PCS or MCS scores (Table 6).

4.3   Ischemic Heart Disease

Table 7 shows that average health plan performance on each of the heart disease-related HEDIS measure was high: 94.6% (SD 8.9%) for Beta Blockers After Heart Attack measure, 79.1% (SD 9.5%) for LDL screenings, and 62.9% (SD 15.4%) for LDL control.

None of the three HEDIS measures for heart disease were significantly associated with changes in individual PCS scores (Table 8) or MCS scores (Table 9). The trend however was positive in five of the six comparisons (all three of the measures with MCS, and beta blocker and LDL control measures with PCS).

4.4   Depression

Health plan performance on measure related to Depression, including the three Antidepressant Medication Management measures and two Follow-Up after Mental Health Hospitalization measures, was generally lower than health plan performance on measures related to diabetes and heart disease. (Table 10) The rates varied from a low of 11.4% for optimal practitioner contact for medication management, to a high of 63.1% for having a follow-up visit within 30 days after hospital discharge.

None of the measures was significantly associated with changes in individual PCS (Table 11) or MCS scores (Table 12). Only the effect of optimal practitioner contact on changes in PCS scores approached statistical significance, however the effect was opposite of what one might expect. A 10 percentage point increase in the proportion of enrollees receiving optimal practitioner contact was associated with a 15 percentage point decrease in the probability of being healthy as measured by PCS change (p=.068).

5.0   DISCUSSION

Of the four disease subgroups, health plan quality performance had a positive impact on longitudinal change in enrollees' health only for those with diabetes, regardless of baseline health differences in plan populations. However, only the composites of diabetes measures achieved statistical significance in their relationship with outcomes as measured by the PCS and MCS. The intermediate outcomes composite (HbA1c control and LDL control) was significantly related to individual improvements in the PCS but was not related to changes in the

MCS. The process composite, on the other hand, was significantly related to individual improvements in the MCS but not the PCS. Additionally, it appears that health plan performance on the intermediate outcomes composite had the largest effect on enrollees who were the healthiest of the enrollees with diabetes. The smaller effect for those persons with intermediate or lower levels of physical or mental functioning (i.e. sicker enrollees) compared to the healthiest enrollees is somewhat surprising. One might expect to see the largest improvements among this group of enrollees. One explanation for this finding may be that higher severity of illness may be such that improved care reflected by HEDIS measures is unable to impact a high, and perhaps more fixed, level of impairment. Sicker enrollees are likely to have multiple co-morbidities and disabilities, and diabetes care alone may not be sufficient to change their health trajectory. Future research should examine this phenomenon more thoroughly and over longer periods of time.

Overall, the observed relationships of the health plan performance HEDIS process and intermediate outcomes composite with changes in PCS and MCS for enrollees with diabetes have face validity. Since diabetes is a condition in which short term interventions, such as better control of HbA1c, might have immediate impacts on day to day functioning. One can postulate that better control the levels of HbA1c, might lead to patients, and especially those with lower overall levels of functional impairment, to experience somewhat improved physical functioning. The process measures, which mostly relate to patient monitoring and not directly to outcomes, show a positive but not significant relationship given that they are one more step removed from impacting outcomes. The stronger, and significant, association of the diabetes process measures with better mental health functioning is somewhat paradoxical, although there was a positive but not significant association of the outcome measures with the MCS. Perhaps individuals who are enrolled in plans that foster better monitoring of their diabetic patients are more satisfied with their care leading to less stress as a result. In other studies we have shown a positive correlation between process measures and patient experience of care in HEDIS.

While in general the direction of the associations were positive, no significant associations between HEDIS plan quality performance and change in PCS or MCS were found in any of the other three disease subgroups – hypertension, ischemic heart disease or depression. While these

results are surprising, there may be various reasons for these findings. For hypertension, the lack of a significant association of plan performance on this measure at one point in time to individual outcomes as measured by the PCS is not surprising. Generally speaking, by itself, hypertension does not cause discomfort or pain in patients over the short term. Indeed, individuals adequately treated to achieve more BP control may, on average, feel worse then those who are untreated. It is only after a stroke or myocardial infarction that BP control might have an impact on patient mental or physical functioning. It is unlikely that there would be enough excess cardiovascular events in the plans where treatment was below average, to see impacts on overall functional status.

For ischemic heart disease, one would expect better individual outcomes in plans that were better at achieving LDL control among its enrollees and better at providing beta blockers after a myocardial infarction. For the measure relating to beta blockers, the mean proportion meeting this criterion was over 94%, so most plans were providing good care and thus there was not significant enough variation in this process measure to detect a significant relationship. Moreover, this measure relates to a relatively small proportion (those with MI's within the year) of the group identified in the HOS as having ischemic heart disease (anyone who had ever had an MI, angina, or been told that they had coronary artery disease. However, there was more variation with LDL control between plans, and one would expect better LDL control to be associated with better physical health outcomes. While we saw a positive association between, a lack of a significant finding may be due to limited power to detect a statistically significant effect.

Finally, for depression care, optimal practitioner contact and effective acute phase and continuation phase treatment, which would reflect careful follow-up, monitoring, and treatment of depressed patients, should logically be related to better MCS outcomes, but these associations were not found. The generally poor performance rates among health plans on these depression treatment process measures among all plans may have contributed to this null finding. Another issue is the rather tenuous relationship between the depression indicator question in the HOS (person reported that they felt sad or depressed most of the time) and the clinical diagnosis of depression. There is ample evidence that depression is substantially under diagnosed, especially

in the Medicare population. The three core HEDIS depression measures do not include a screening measure, and focus only on follow-up of treatment of persons diagnosed with depression and placed on medication. By contrast, the self identification used in the HOS is likely to include a much larger group of enrollees than the denominator for the HEDIS measure. Thus any effect of relatively small differences in a measures with low overall performance, is likely undetectable in this situation. Finding a significant impact of the measures related to follow-up care following a hospitalization is even more unlikely due to the small number of enrollees that are hospitalized for depression during the year. Since most enrollees with diagnosed depression, let alone the even larger number with self-identified depression, are unlikely to need inpatient treatment for this condition, the majorities of individual enrollees are unlikely to have received this type of care and thus would not be expected to have improvements in outcome related to this measure.

Limitations

There are several limitations to this study. First, there is plan attrition between each baseline and follow-up survey period. Plans that dropped out of the MA program were not included in the study database for the affected HOS cohorts; hence the study results are based on the more stable plans that remain in the program for each HOS cycle. In addition, there is beneficiary attrition during the course of the HOS administration due to baseline and follow-up non-response, as well as voluntary and involuntary dis-enrollment between surveys. These factors could introduce bias and may limit our ability to generalize the results of the study to populations with shorter periods of continuous enrollment. However, HOS response rates are generally high, with around a 60% response rate in baseline years that the survey is fielded, and around an 80% response rate in the two-year follow-up survey. Any associations we find should describe the situation in stable Medicare managed care populations as well as fee-for-service populations in stable patient-provider relationships.

A second limitation is the reliance on self-reported disease status for the identification of enrollees with the four diseases. There is the possibility of misidentification, as some enrollees who identify themselves as having the disease, may not meet the clinical definition of the disease and even more importantly, may not fit the group specified to be included in the denominator of

Table 1   Prevalence of Disease-Specific Responses in HOS Cohort 4

(Relative to Total Eligible Responses, n=48,521)

Disease Subgroup	Cohort 4	
	Number of Disease-Specific Responses	**Disease-Specific Responses as % of Total Eligible Responses (n=48,521)**
Diabetes	8,184	16.9%
Hypertension	27,206	56.1%
Ischemic Heart Disease	9,125	18.8%
Depression	4,911	10.1%

Table 2   Relevant HEDIS Measures for Four Disease Subgroups

Subgroup	Measure Name
Diabetes	<u>Comprehensive Diabetes Control: Intermediate Outcomes</u> % of patients with Hemoglobin A1c Control - < 9.5 % of patients with LDL Cholesterol Control - < 130 *Comprehensive Diabetes Control: Process* % of patients with Hemoglobin A1C Testing % of patients with LDL Cholesterol Screening % of patients with Eye Exam Performed % of patients with Nephropathy Monitored
Hypertension	<u>Controlling High Blood Pressure</u> % of individuals with diagnosed hypertension with BP less than 140/90
Ischemic Heart Disease	<u>Beta-Blockers After AMI</u> % of AMI patients with beta-blocker prescriptions within 7 days of discharge <u>Cholesterol Screening After Acute Cardiovascular Event</u> % of patients with cholesterol screening following AMI, CABG, or PTCA <u>Cholesterol Control After Acute Cardiovascular Event</u> % of patients with LDL Cholesterol Control <130 following AMI, CABG, or PTCA
Depression	<u>Antidepressant Medication Management</u> % of depressed patients with at least 3 follow up visits during the 3 months post-diagnosis % of depressed patients continuing on antidepressants for 3 months post-diagnosis (acute phase treatment) % of depressed patients continuing on antidepressants for 6 months post-diagnosis (continuation phase treatment) And <u>7 and 30-Day Outpatient Follow-up After Hospitalization for Mental Illness</u>

Table 3   Condition-Specific Symptom Severity Definitions

Symptom Severity Scale	Definition
*Diabetic Neuropathy**	
With	Any of four symptoms (numbness/loss of feeling in feet, or tingling/burning feet, or decreased feeling of hot/cold with feet, or sores not healing on feet) reported all or most of the time.
Without	Any of these four symptoms reported some, a little, or none of the time.
*Shortness of Breath***	
4 Severe	Short of breath all or most of the time when lying down flat, or sitting, or resting
3 Moderate	Short of breath all or most of the time when walking less than 1 block, or when climbing 1 flight of stairs
2 Mild	Short of breath some or a little of the time when walking les than 1 block, or sitting, or resting, or when climbing 1 flight of stairs, or lying down flat
1 Asymptomatic	Short of breath none of the time when lying down flat and sitting/resting, and walking less than 1 block, and climbing 1 flight of stairs
*Chest Pain***	
4 Severe	Chest pain or pressure all or most of the time when exercising and resting
3 Moderate	Chest pain or pressure all or most of the time when exercising, and chest pain/pressure some or a little of the time when resting
2 Mild	Chest pain/pressure some or a little of the time when exercising
1 Asymptomatic	Chest pain none of the time when exercising and resting
*Used in Model for Diabetes	
**Used in Model for Ischemic Heart Disease	

Table 4   Health Plan Performance Rates for HEDIS Diabetes Measures 2002

(N=152 plans) [1]

	Mean	Standard Deviation	Minimum	Maximum
Process Composite[2]	76.4%	8.4%	41.3%	90.1%
HbA1c Testing	87.7%	8.7%	23.2%	97.1%
Eye Exams	71.0%	10.9%	28.2%	92.8%
LDL Screenings	89.3%	7.1%	41.0%	98.5%
Monitoring Nephropathy	57.7%	15.0%	26.9%	93.2%
Intermediate Outcomes Composite[3]	70.9%	11.3%	14.9%	88.6%
HbA1c Control[4]	77.6%	12.8%	0.3%	92.7%
LDL Control	64.1%	11.1%	26.4%	85.9%

[1]Each rate represents the proportion of eligible health plan enrollees who met the indicators.

[2]Note: The process composite represents the average proportion of enrollees with diabetes in each plan with HbA1c testing, eye exams, LDL screenings, and monitoring of nephropathy.

[3] The intermediate outcomes composite represents the average proportion of enrollees with diabetes in each plan with HbA1c control and LDL control.

[4]Because the HEDIS measure evaluates poor control of HbA1c (>9.5), this measure was inverted before calculating the composite.

Table 5   Impact of Health Plan Process and Intermediate Outcomes Composite Scores on Changes in PCS and MCS Score or Enrollees with Diabetes (N=8,184)[1]

	Physical Component Score (PCS)		Mental Component Score (MCS)	
	Process Composite[2] β – estimate(p-value)	Intermediate Outcomes Composite β -estimate (p-value)	Process Composite β -estimate (p-value)	Intermediate Outcomes Composite β - estimate (p-value)
HEDIS Composite	0.041 (0.362)	0.082 (0.014)	0.108 (0.003)	0.039 (0.162)
HEDIS Composite*MCS/PCS Rank 1[3,4]	-0.043 (0.485)	-0.076 (0.093)	-0.062 (0.217)	0.008 (0.842)
HEDIS Composite*MCS/PCS Rank 2	-0.055 (0.379)	-0.117 (0.012)	-0.063 (0.211)	-0.020 (0.602)

[1] Based on hierarchical linear regressions controlling for age, gender, education level, race, marital status, home ownership, number of chronic conditions, diabetes symptom severity, presence of depressed mood, and baseline PCS and MCS.

[2]The process composite represents the average proportion of enrollees with diabetes in each plan with HbA1c testing, eye exams, LDL screenings, and monitoring of nephropathy. The intermediate outcomes composite represents the average proportion of enrollees with diabetes in each plan with good HbA1c control and LDL control.

[3]PCS Rank 1 and PCS Rank 2 are indicator variables based on baseline ranking of the physical component summary of the MOS SF-36. MCS Rank 1 and MCS Rank 2 are indicator variables based on baseline ranking of the mental component summary of the MOS SF-36. An individual with a baseline ranking in the bottom tertile (sickest) receives a value of 1 for PCS or MCS Rank, and an individual with a ranking in the middle tertile receives a value of 2 for PCS or MCS Rank.

[4]HEDIS Score*PCS Rank 1 is the HEDIS score interacted with individual enrollee baseline PCS score in the bottom third of PCS scores. HEDIS Score*PCS Rank 2 is the HEDIS score interacted with individual enrollee baseline PCS score in the middle third of PCS scores.

Table 6   Impact of Health Plan Performance on HEDIS Controlling High Blood Pressure Measure on Changes in PCS and MCS Score for Enrollees with Hypertension (N=27,206)[1]

	Physical Component Score, PCS β –estimate (p-value)	Mental Component Score, MCS β -estimate (p-value)
HEDIS Score	0.049 (0.109)	0.032 (0.196)
HEDIS Score*PCS Rank 1[2,3]	-0.041 (0.339)	-0.038 (0.302)
HEDIS Score*PCS Rank 2	-0.004 (0.919)	-0.029 (0.422)

[1] Based on hierarchical linear regressions controlling for age, number of chronic conditions, and baseline PCS and MCS.

[2]PCS Rank 1 and PCS Rank 2 are indicator variables based on baseline ranking of the physical component summary of the MOS SF-36. MCS Rank 1 and MCS Rank 2 are indicator variables based on baseline ranking of the mental component summary of the MOS SF-36. An individual with a baseline ranking in the bottom tertile (sickest) receives a value of 1 for PCS or MCS Rank, and an individual with a ranking in the middle tertile receives a value of 2 for PCS or MCS Rank.

[3]HEDIS Score*PCS Rank 1 is the HEDIS score interacted with individual enrollee baseline PCS score in the bottom third of PCS scores. HEDIS Score*PCS Rank 2 is the HEDIS score interacted with individual enrollee baseline PCS score in the middle third of PCS scores.

Table 7   Health Plan Performance Rates for HEDIS Heart Disease Measures 2002

(N=152 plans)*

	Mean	Std. Dev.	Minimum	Maximum
Beta Blocker After Heart Attack	94.6%	8.9%	31.4%	100.0%
LDL Screening	79.1%	9.5%	35.1%	98.0%
LDL Control <130	62.9%	15.4%	0.0%	96.0%

*These measures are based on HEDIS health plan reporting; each rate represents the proportion

of health plan enrollees who met the indicators.

Table 8   Impact of Health Plan Performance on HEDIS Heart Disease Measures on Changes in PCS Score for Enrollees with Ischemic Heart Disease (N=9,125)[1]

	Beta Blocker after Heart Attack β –estimate (p-value)	LDL Cholesterol Screenings β –estimate (p-value)	LDL Cholesterol Control <130mg/dl β –estimate (p-value)
HEDIS Score	0.022 (0.626)	-0.010 (0.803)	0.030 (0.213)
HEDIS Score*PCS Rank 1[2,3]	-0.009 (0.891)	0.019 (0.734)	-0.026 (0.459)
HEDIS Score*PCS Rank 2	0.036 (0.583)	0.010 (0.855)	-0.013 (0.701)

[1] Based on hierarchical linear regressions controlling for age, number of chronic conditions, presence of depressed mood, ischemic heart disease symptom severity (chest pain and shortness of breath), and baseline PCS and MCS.

[2]PCS Rank 1 and PCS Rank 2 are indicator variables based on baseline ranking of the physical component summary of the MOS SF-36. MCS Rank 1 and MCS Rank 2 are indicator variables based on baseline ranking of the mental component summary of the MOS SF-36. An individual with a baseline ranking in the bottom tertile (sickest) receives a value of 1 for PCS or MCS Rank, and an individual with a ranking in the middle tertile receives a value of 2 for PCS or MCS Rank.

[3]HEDIS Score*PCS Rank 1 is the HEDIS score interacted with individual enrollee baseline PCS score in the bottom third of PCS scores. HEDIS Score*PCS Rank 2 is the HEDIS score interacted with individual enrollee baseline PCS score in the middle third of PCS scores.

Table 9   Impact of Health Plan Performance on HEDIS Heart Disease Measures on Changes in MCS Score for Enrollees with Ischemic Heart Disease (N=9,125)[1]

	Beta Blocker After Heart Attack β –estimate (p-value)	LDL Cholesterol Screenings β –estimate (p-value)	LDL Cholesterol Control <130 β –estimate (p-value)
HEDIS Score	0.037 (0.245)	0.020 (0.533)	0.021 (0.267)
HEDIS Score*MCS Rank 1[2,3]	-0.096 (0.811)	0.058 (0.212)	-0.034 (0.233)
HEDIS Score*MCS Rank 2	-0.019 (0.686)	-0.008 (0.850)	-0.003 (0.928)

[1] Based on hierarchical linear regressions controlling for age, number of chronic conditions, presence of depressed mood, ischemic heart disease symptom severity (chest pain and shortness of breath), and baseline PCS and MCS.

[2]PCS Rank 1 and PCS Rank 2 are indicator variables based on baseline ranking of the physical component summary of the MOS SF-36. MCS Rank 1 and MCS Rank 2 are indicator variables based on baseline ranking of the mental component summary of the MOS SF-36. An individual with a baseline ranking in the bottom tertile (sickest) receives a value of 1 for PCS or MCS Rank, and an individual with a ranking in the middle tertile receives a value of 2 for PCS or MCS Rank.

[3]HEDIS Score* PCS Rank 1 is the HEDIS score interacted with individual enrollee baseline PCS score in the bottom third of PCS scores. HEDIS Score*PCS Rank 2 is the HEDIS score interacted with individual enrollee baseline PCS score in the middle third of PCS scores.

Table 10   Health Plan Performance Rates for HEDIS Depression-Related Measures 2002

(N=152 plans)*

	Mean	Std. Dev.	Minimum	Maximum
Optimal Practitioner Contact Post-Diagnosis	11.4%	7.0%	2.1%	38.4%
Effective 3-Month (Acute) Treatment Post-Diagnosis	53.0%	10.4%	28.2%	77.0%
Effective 6-Month (Continuation) Treatment Post-Diagnosis	38.6%	11.5%	11.1%	67.2%
Follow-up 7 Days Post-Hospitalization	41.4%	18.4%	5.9%	86.7%
Follow-up 30 Days Post-Hospitalization	63.1%	15.8%	22.2%	91.1%

*These measures are based on HEDIS health plan reporting; each rate represents the proportion

of health plan enrollees who met the indicators.

Table 11   Impact of Health Plan Performance on HEDIS Depression Related measures on Changes in PCS Score for Enrollees with Depression (N=4,911) [1]

	Optimal Practitioner Contact β –estimate (p-value)	Effective 3-Month (Acute) Treatment β –estimate (p-value)	Effective 6-Month (Continuation) Treatment β –estimate (p-value)	7 Day Follow-Up β –estimate (p-value)	30 Day Follow-up β –estimate (p-value)
HEDIS Score	-0.153 (0.068)	-0.030 (0.598)	0.045 (0.395)	-0.010 (0.757)	-0.028 (0.455)
HEDIS Score*PCS Rank 1[2,3]	0.147 (0.224)	0.080 (0.339)	-0.009 (0.910)	0.016 (0.718)	0.040 (0.437)
HEDIS Score*PCS Rank 2	0.204 (0.099)	0.059 (0.464)	0.006 (0.930)	-0.028 (0.528)	0.001 (0.989)

[1] Based on hierarchical linear regressions controlling for age, number of chronic conditions, and baseline PCS and MCS.

[2] PCS Rank 1 and PCS Rank 2 are indicator variables based on baseline ranking of the physical component summary of the MOS SF-36. MCS Rank 1 and MCS Rank 2 are indicator variables based on baseline ranking of the mental component summary of the MOS SF-36. An individual with a baseline ranking in the bottom tertile (sickest) receives a value of 1 for PCS or MCS Rank, and an individual with a ranking in the middle tertile receives a value of 2 for PCS or MCS Rank.

[3] HEDIS Score*PCS Rank 1 is the HEDIS score interacted with individual enrollee baseline PCS score in the bottom third of PCS scores. HEDIS Score*PCS Rank 2 is the HEDIS score interacted with individual enrollee baseline PCS score in the middle third of PCS scores.
NOTE: The model controlled for age and number of chronic conditions.

Table 12   Impact of Health Plan Performance on HEDIS Depression Related measures on Changes in MCS Score for Enrollees with Depression (N=4,911) [1]

	Optimal Practitioner Contact β –estimate (p-value)	Effective 3-Month (Acute) Treatment β –estimate (p-value)	Effective 6-Month (Continuation) Treatment β –estimate (p-value)	7 Day Follow-Up β –estimate (p-value)	30 Day Follow-up β –estimate (p-value)
HEDIS Score	0.126 (0.103)	-0.002 (0.974)	-0.014 (0.763)	-0.022 (0.422)	-0.023 (0.471)
HEDIS Score*MCS Rank 1[2,3]	-0.031 (0.778)	-0.008 (0.917)	0.006 (0.929)	-0.025 (0.543)	-0.005 (0.917)
HEDIS Score*MCS Rank 2	-0.179 (0.119)	0.040 (0.596)	0.034 (0.613)	-0.001 (0.981)	-0.015 (0.744)

[1] Based on hierarchical linear regressions controlling for age, number of chronic conditions, and baseline PCS and MCS.

[2] PCS Rank 1 and PCS Rank 2 are indicator variables based on baseline ranking of the physical component summary of the MOS SF-36. MCS Rank 1 and MCS Rank 2 are indicator variables based on baseline ranking of the mental component summary of the MOS SF-36. An individual with a baseline ranking in the bottom tertile (sickest) receives a value of 1 for PCS or MCS Rank, and an individual with a ranking in the middle tertile receives a value of 2 for PCS or MCS Rank.

[3] HEDIS Score*PCS Rank 1 is the HEDIS score interacted with individual enrollee baseline PCS score in the bottom third of PCS scores. HEDIS Score*PCS Rank 2 is the HEDIS score interacted with individual enrollee baseline PCS score in the middle third of PCS scores.
NOTE: The model controlled for age and number of chronic conditions.

REFERENCES

[1] HEDIS is a registered trademark of the National Committee for Quality Assurance.

[2] Eisenberg, J. (February 26, 1998). *Assessing Health Care Quality.* Statement on behalf of the Agency for Health Care Policy and Research before the Subcommittee on Health, Committee on Ways and Means, U.S. House of Representatives, Washington, D.C.: U.S. Government Printing Office, Serial No. 105-94.

[3] McGlynn, EA (2003). Selecting common measures of quality and system performance. *Medical Care*, 41(1, supplement), I39-47.

[4] Palmer, RH (1995). Securing health care quality for Medicare. *Health Affairs*, 14(4), 89-100.

[5] For example, HbA1c screening and control in diabetics, cholesterol screening and control post MI.

[6] Eddy, D. M. (1998). Performance measurement: Problems and solutions. *Health Affairs, 17 (4)*, 7-25.

[7] Palmer RH, (1998). Using health outcomes data to compare plans, networks and providers. *International Journal for Quality in Health Care,* 10(6):477-483.

[8] National Academy Press. (1999). *Measuring the Quality of Health Care,* Washington, D.C.

[9] Agency for Healthcare Research and Quality (March, 2001). *AHRQ profile: Quality research for Quality Health Care*. AHRQ publication No. 00-P005. Rockville, MD: Author. Available on the world wide web at: http://www.ahrq.gov/about/profile.htm

[10] Eisenberg, J. (February 26, 1998). *Assessing Health Care Quality*. Statement on behalf of the Agency for Health Care Policy and Research before the Subcommittee on Health, Committee on Ways and Means, U.S. House of Representatives, Washington, D.C.: U.S. Government Printing Office, Serial No. 105-94.

[11] McGlynn, EA (2003). Selecting common measures of quality and system performance. *Medical Care*, 41(1, supplement), I39-47.

[12] Palmer, RH (1995). Securing health care quality for Medicare. *Health Affairs*, 14(4), 89-100.

the HEDIS measures. As we have noted, the highest correspondence between self report, and the group included in the HEDIS measure denominator, is for those persons with diabetes. The denominator inclusion criteria for measures related to beta blocker use in the period immediately after an MI, or those for cholesterol screening or control in patients after an MI or major coronary artery procedure are much more restrictive than the self definitions used in the HOS. For enrollees with depression, there may be an even greater disparity in denominator definition of the measure and self-report since the depression item in HOS might be described as asking about depressed mood rather than clinical depression. While there is evidence that older adults can reliably report their chronic conditions and that the positive predictive value of such reports is high, the problem of under diagnosis of depression is a major confounding factor.[44] [45] A countervailing issues is that while the HEDIS quality measures do not capture all aspects of plan services that may influence outcomes for these patients, they do represent key evidence-based aspects of care that we have shown in other studies to have significant correlations with overall plan quality in each clinical area.

Conclusions

Despite these limitations, this study represents the first attempt to directly link plan performance on HEDIS quality measures to changes in individual enrollee outcomes. In summary, it appears that plan performance on HEDIS measures for diabetes, which are the most closely linked in terms of HOS and HEDIS measure denominators, appear to be related to individual enrollee outcomes, suggesting that the HEDIS measures are very likely to be valid measures of how well a health plan is providing care to its enrollees. Further studies using directly linked patient level data and identical definitions of the populations included on both HEDIS clinical measures and the tool used to assess physical and mental functioning are clearly indicated.

[13] Lied TR, Sheingold SH (2001). Relationships among performance measures for Medicare managed care plans. *Health Care Financing Review*, 22 (3), 23-33.

[14] Cooper JK, et al, (2001). Health outcomes: New quality measure for Medicare. *International Journal for Quality in Health Care*, 13(1):9-16.

[15] HEDIS measures definitions and sampling details can be found *in HEDIS Volume 2 - Technical Specifications* or on the HEDIS section of the NCQA Website (http://www.ncqa.org/programs/hedis/).

[16] HEDIS measure definitions and sampling details can be found *in HEDIS Volume 2 - Technical Specifications* or on the HEDIS section of the NCQA Website (http://www.ncqa.org/programs/hedis/).

[17] Details on the HOS survey instrument and administration procedures may be found in *HEDIS Volume 6 – Specifications for the Medicare Health Outcomes Survey* or on the HOS Website (http://www.hosonline.org).

[18] MOS SF-36® is a registered trademark of the Medical Outcomes Trust.

[19] Ware JE, et.al. (1999). Health related quality of life in chronic Hepatitis C: Impact of disease and treatment response. *Hepatology*, 30:550-5.

[20] Garratt AM, et.al. (1994). SF-36 Health Survey Questionnaire II: Responsiveness to changes in health status in four common clinical conditions. *Quality in Health Care*, 3:186-92.

[21] Boult C, et.al. (2001). A randomized clinical trial of outpatient geriatric evaluation and Management. Journal *American Geriatric Society*, 49:351-9.

[22] Garratt AM, et.al. (1993). The SF-36 Health Survey Questionnaire: An outcome measure suitable for routine use within the NHS? *British Medical Journal*, 306:1440-4.

[23] Ahroni JH, Boyko EJ, (2000). Responsiveness of the SF-36 among veterans with Diabetes Mellitus. *Journal of Diabetes and Its Complications*, 14:31-39.

[24] Lau C, Qureshi AK, Scott SG (2004). Association between glycemic control and quality of life in diabetes mellitus. *Journal of Postgraduate Medicine*, 50 (3), 189-94.

[25] Helstad CP, Wiegmann SM, Camponeschi J, Udlis S, Meredith M (2004). The Wisconsin Medical Society diabetes study. *Wisconsin Medical Journal*, 103 (3), 71-8.

[26] Ibrahim IA, et al, (2002). Measuring outcomes of type 2 Diabetes disease management Program in an HMO setting. *Southern Medical Journal*, 95(1):78-87.

[27] Clancy CM, and Bierman AS, (2000). Quality and outcomes of care for older women with chronic disease. *Women's Health Issues,* 10(4):178-190.

[28] Tsai JC, Yang HY, Wang WH, Hsieh MH, Chen PT, Kao CC, Kao PF, Wang CH, Chan P (2004). The beneficial effect of regular endurance exercise training on blood pressure and quality of life in patients with hypertension. *Clinical and Experimental Hypertension,* 26 (3), 255-65.

[29] Charlier L, Dutrannois J, and Kaufman L, (1997). The SF-36 Questionnaire: A convenient way to assess quality of life in angina pectoris patients, *Acta Cardiologica,* 52(3):247-60.

[30] Fihn SD, McDonell MB, Diehr P, Anderson SM, Bradley KA, Au DH, Spertus JA, Burman M, Reiber GE, Kiefe CI, Cody M, Sanders KM, Whooley MA, Rosenfeld K, Baczek LA, Sauvigne A (2004). Effects of sustained audit/feedback on self-reported health status of primary care patients. *American Journal of Medicine,* 116 (4), 241-8.

[31] Ohldin A, Young B, Derleth A, McDonell M, Diehr P, Kiefe C, Fihn S (2004). Ethnic differences in satisfaction and quality of life in veterans with ischemic heart disease. J National Medical Association, 96 (6), 799-808.

[32] Coulehan JL, et al, (1997). Treating depressed primary care patients improves their physical, mental, and social functioning, *Archives of Internal Medicine,* 157(10):1113-20.

[33] Fifer SK, et al, (2003). Functional status and somatization as predictors of medical offset in anxious and depressed patients, *Value in Health*, 6(1):40-50.

[34] Williams JW Jr, Barrett J, Oxman T, Frank E, Katon W, Sullivan M, Cornell J, Sengupta A (2000). Treatment of dysthymia and minor depression in primary care: A randomized controlled trial in older adults. *Journal of the American Medical Association,* 284 (12), 1519-26.

[35] Stein MB, Barrett-Connor E (2002). Quality of life in older adults receiving medications for anxiety, depression, or insomnia: findings from a community-based study. *American Journal of Geriatric Psychiatry*, 10 (5), 568-74.

[36] Details on the HOS survey instrument and administration procedures may be found in *HEDIS Volume 6 – Specifications for the Medicare Health Outcomes Survey* or on the HOS Website (http://www.hosonline.org).

[37] Ware JE, Kosinski M, Keller SD (1994). *SF-36 Physical and Mental Health Summary Scales: A User's Manual*. Boston, MA: The Health Institute.

[38] Ware JE, and Kosinski M (2001). Interpreting SF-36 summary health measures: A response, *Quality of Life Research*, 10:405-13.

[39] The eight scale scores were calculated for cohorts 1-3 only due to a change in the scoring methodology. The summary measures (PCS and MCS) and the individual questions continue to be available for analysis of subsequent cohorts.

[40] Diehr P and Patrick DL (2003). Trajectories of health for older adults over time: Accounting fully for death. *Annals of Internal Medicine*. 139:416-420.

[41] Diehr P, Patrick DL, Spertus J, Kiefe CI, McDonell M, and Fihn SD (2001). Transforming self-rated health and the SF-36 scales to include death and improve interpretability. *Medical Care,* 2001 Jul: 39:670-680.

[42] Diehr P, Patrick DL, McDonell MB and Fihn SD (2003). Accounting for Deaths in Longitudinal Studies Using the SF-36: The Performance of the Physical Component Scale of the Short Form 36-Item Health Survey and the PCTD." *Medical Care,* 41: 1065-1973.

[43] Raudenbush, SW and Bryk AS, (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods, 2nd Ed*. Thousand Oaks, CA: Sage Publications Inc.

[44] Silliman RA, and Lash TL, (1999). Comparison of interview-based and medical-record based indices of co-morbidity among breast cancer patients, *Medical Care*, 37(4):339-49.

[45] Miller D.R.,Rogers W.H., Spiro A, Kazis L.E, (2003). Evaluation of Disease Status based on Patient Self-Report in the Medicare Health Outcomes Survey. Available online at www.hosonline.org.