



MEDICARE HEALTH OUTCOMES SURVEY

**REPORT ON THE COMPARISON OF SENIORS IN
PROGRAM OF ALL-INCLUSIVE CARE FOR THE
ELDERLY (PACE) AND MEDICARE ADVANTAGE
PLANS**

Cohorts II, III, IV and V Baseline

FINAL REPORT

**PREPARED BY HEALTH SERVICES ADVISORY GROUP
OCTOBER 26, 2005**



TABLE OF CONTENTS

	<u>PAGE</u>
EXECUTIVE SUMMARY	1
1. INTRODUCTION	3
2. METHODOLOGY	5
3. RESULTS	9
4. DISCUSSION	13
5. CONCLUSIONS	15
6. REFERENCES	16
7. APPENDIX	18

LIST OF FIGURES AND TABLES

	<u>PAGE</u>
FIGURE 1 <i>COHORT V BASELINE</i> SF-36 SUMMARY SCORES	2
TABLE 1 <i>COHORT V BASELINE</i> DEMOGRAPHIC SUMMARY FOR THE PACE SAMPLE AND THE MA SAMPLE.....	18
TABLE 2 <i>COHORT V BASELINE</i> HEALTH STATUS SUMMARY FOR THE PACE SAMPLE AND THE MA SAMPLE.....	19
TABLE 3 <i>COHORT V BASELINE</i> PHYSICAL COMPONENT SUMMARY (PCS) SCORES FOR THE PACE SAMPLE AND THE MA SAMPLE.....	20
TABLE 4 <i>COHORT V BASELINE</i> MENTAL COMPONENT SUMMARY (MCS) SCORES FOR THE PACE SAMPLE AND THE MA SAMPLE	21
TABLE 5 <i>COHORT V BASELINE</i> PHYSICAL COMPONENT SUMMARY (PCS) SCORES FOR THE PACE SAMPLE AND THE MA AGE & ADL MATCHED SAMPLE	22
TABLE 6 <i>COHORT V BASELINE</i> MENTAL COMPONENT SUMMARY (MCS) SCORES FOR THE PACE SAMPLE AND THE MA AGE & ADL MATCHED SAMPLE	23
TABLE 7 <i>COHORTS II, III, IV AND V BASELINE</i> SF-36 SUMMARY SCORES	24

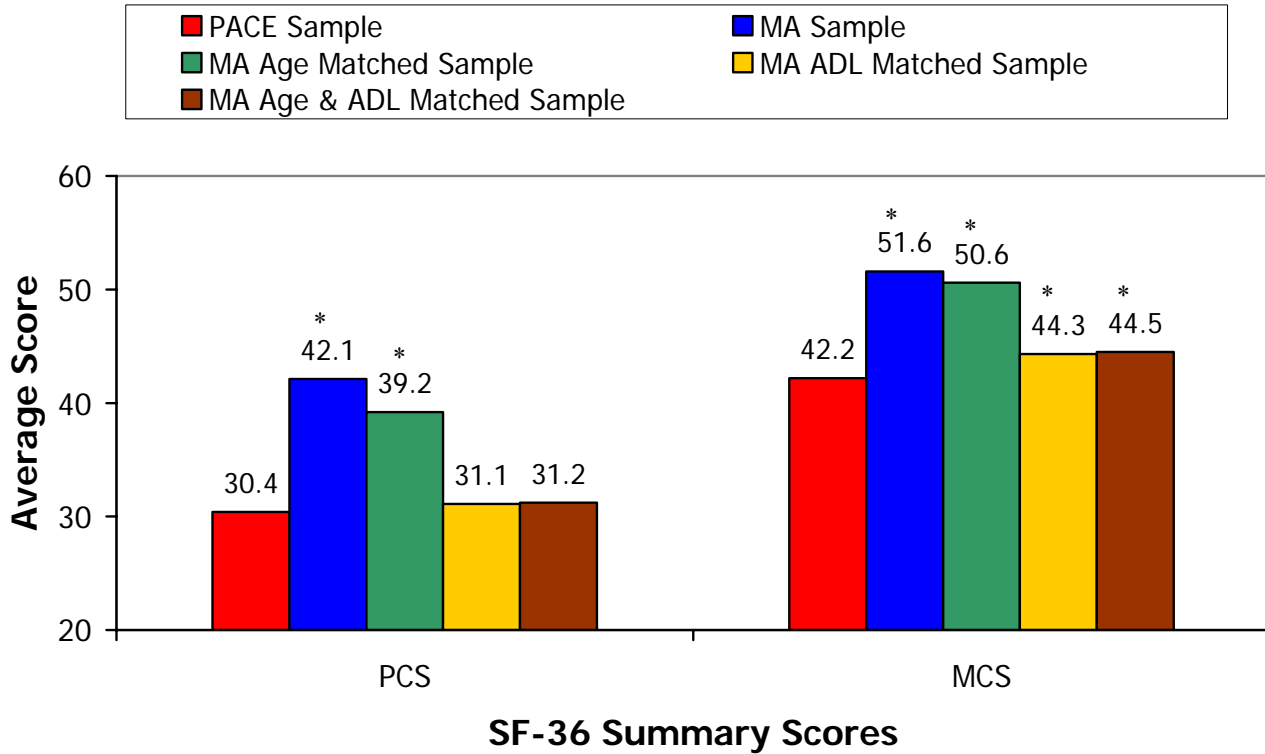
EXECUTIVE SUMMARY

Health Services Advisory Group compared the health status of beneficiaries enrolled in Program of All-Inclusive Care for the Elderly (PACE) plans with the health status of beneficiaries enrolled in Medicare managed care plans who participated in *Cohorts II, III, IV and V Baseline* (1999 to 2002) of the Medicare Health Outcomes Survey (HOS).

This report presents demographic information, self-reported health status, physical and mental health status as measured by the SF-36 health survey, prevalence of chronic conditions, and ratings of activities of daily living (ADLs). Data from the HOS provide an opportunity to compare physical and mental health status, age group distributions, and functional limitations between the Medicare respondents enrolled in PACE and managed care plans.

Descriptive analyses indicate that the Medicare managed care respondents were younger and had significantly better physical and mental health summary measures than the PACE respondents. When managed care respondents were randomly matched to PACE respondents by age group and functional limitations, physical and mental health summary measures were no longer significantly different. These results suggest that comparisons between beneficiaries of PACE and managed care plans should be controlled for both age and functional limitations.

FIGURE 1
COHORT V BASELINE SF-36 SUMMARY SCORES



*Significantly different by *t* test from the corresponding PACE sample mean at the $p < 0.01$ level.

Source: Cohort V Baseline from the Medicare Health Outcomes Survey Database

1

INTRODUCTION

MEDICARE MANAGED CARE

In 1982, Congress mandated the provision of managed care plan options to Medicare beneficiaries through the U.S. Department of Health and Human Services (DHHS). In 1997, the Balanced Budget Act (BBA) established a new Part C of the Medicare program (DHHS, 2003). Known originally as Medicare+Choice (M+C) plans, Medicare managed care plans became known as Medicare Advantage (MA) plans in 2004. The Medicare managed care program provides eligible individuals with the option of receiving Medicare benefits through a managed care setting in lieu of the traditional Medicare fee-for-service (FFS) program. In addition to the original Medicare coverage, which provides hospital insurance and supplementary medical insurance, Medicare managed care plans may provide additional benefits such as coordination of care, reduction of out-of-pocket expenses, prescription drug coverage, dental care, and routine physical and vision services. Of a total of 40 million Americans in Medicare, about 5.6 million (15 percent) have chosen to be in an MA plan (Centers for Medicare & Medicaid Services [CMS], 2001).

PACE PROGRAM

The BBA of 1997 authorized coverage of PACE programs under the Medicare program, and the establishment of PACE as a state option under Medicaid (DHHS, 1999). PACE plans are organizations that provide enhanced services to individuals who meet the following eligibility criteria: 1) are age 55 or older; 2) are certified by their state as needing nursing home care; 3) are able to live safely in the community upon enrollment; and 4) live in a PACE geographical catchment area (CMS, 2002). PACE is only available in states that have chosen to offer the program under the Medicaid state plan. Many of the PACE plan beneficiaries are dual-eligible, receiving both Medicare and Medicaid assistance.

The PACE program emphasizes improving social and psychological well being, physical and mental health status, functional independence, and quality of life, with the overall goal of keeping beneficiaries in their homes and communities. The core services include adult day health care and management of all health services by an interdisciplinary team. The team directs the provision of any needed medical, supportive, and rehabilitative services, including nursing home care as necessary. According to the National PACE Association (NPA, 2002),¹ approximately 7,000 individuals were enrolled in PACE by the end of 2000. Currently there are 25 programs in 14 states that offer PACE services (CMS, 2005).

¹ Membership included 25 PACE programs and 11 pre-PACE programs. The latter operated under Medicaid contracts which cover only long-term care services and are billed under Medicare FFS.

Functional status reflects an individual's ability to carry out Activities of Daily Living (ADLs) and to participate in society. Functional disability, mental health impairment, and the lack of social support have been associated with long-term care institutionalization (Eng, Pedulla, Eleazer, McCann, & Fox, 1997). Although PACE serves a population at high risk for institutional care, only 7.6 percent of PACE beneficiaries are permanent nursing home residents (NPA, 2002).

Programs such as PACE provide a comprehensive alternative to long-term care, incurring lower costs than with nursing home care (Bodenheimer, 1999). PACE sites save money by keeping beneficiaries out of the hospital. PACE beneficiaries tend to be sicker and more disabled than traditional Medicare FFS beneficiaries; however, the number of hospital days for PACE beneficiaries in 1998 was comparable with the number of hospital days in 1997 for Medicare FFS beneficiaries (Weiland, Lamb, Sutton, Boland, Clark, Friedman et al., 2000). An evaluation of PACE by Abt Associates found that the quality of life and functional status of PACE beneficiaries were better than in a similar non-PACE population (Chatterji, Burstein, Kidder, & White, 1998). The evaluation found that PACE beneficiaries with high levels of ADL limitations experienced the most marked decreases in hospital utilization and nursing home days compared to the non-PACE beneficiaries with fewer limitations.

Data from the Medicare HOS provide an opportunity to compare the physical and mental health status, age group distributions, and functional limitations between the Medicare respondents enrolled in PACE and those enrolled in managed care plans.

2

METHODOLOGY

BACKGROUND

The Medicare HOS is a longitudinal survey that assesses the physical and mental health status of Medicare beneficiaries enrolled in managed care in the United States. Beginning in 1998 and continuing annually, a new baseline cohort is created from a randomly selected sample of 1,000 Medicare managed care enrollees from each applicable Medicare contract market area. In plans with fewer than 1,000 Medicare beneficiaries, the sample includes the entire enrolled Medicare population that meets the inclusion criteria. Medicare beneficiaries who are continuously enrolled in the health plans for at least six months are eligible for sampling. Each baseline cohort is surveyed again two years later (the follow-up cohort).

Since the individual PACE plan memberships were less than 1,000 for *Cohorts II* through *V*, all beneficiaries of the PACE plans who met the inclusion criteria were eligible for the survey. In addition, the six-month requirement was waived for beneficiaries of the PACE program. HOS data from *Cohorts II, III, IV, and V Baseline*, which were fielded from 1999 to 2002, were utilized for this analysis. *Cohort I* was not included in this analysis because PACE plans did not participate in this cohort.

The data collection protocol includes a combination of mail and telephone surveys. Multiple mailings, standardized telephone interviews, interviewer training, and methods for maximizing response rates are well-established in the HEDIS^{®2} specifications (National Committee for Quality Assurance [NCQA], 2003).

MEDICARE HOS INSTRUMENT

The Medicare HOS instrument includes the SF-36 health survey, which is a widely used multi-purpose, short-form health survey. Reliability and validity of the SF-36 have been well established (McHorney, Ware, Lu, & Sherbourne, 1994). The SF-36 yields an eight-scale profile of scores and is a generic measure, as opposed to one that targets a specific age, disease, or treatment group. The eight scales form two distinct higher-ordered clusters that are the basis for scoring the Physical Component Summary (PCS) measure and Mental Component Summary (MCS) measure. For this analysis, the SF-36 individual scale scores, as well as the PCS and MCS scores, have been normed to the values for the 1998 general U.S. population, so that a score of 50 represents the national average for a given scale or summary score, with a standard deviation of 10. Higher scores on the SF-36 measures represent better physical and/or mental health status. In addition to the SF-36, demographic data; activities of daily living (ADLs); 13 chronic conditions; 3 depression-screening questions; current smoking history; 12 clinical

² HEDIS[®] is a registered trademark of the National Committee for Quality Assurance (NCQA).

symptoms relating to chest pain, shortness of breath (SOB), and peripheral neuropathy; and 6 medical problems are assessed.

ANALYSIS CATEGORIES

To facilitate analysis, responses to several of the items have been consolidated into the following categories:

The number of ADL impairments was added together for six ADLs, and ranged from 0 for no impairments to 6 for impairment on all ADLs. The question prefacing all ADL items is: *“Because of a health or physical problem, do you have any difficulty doing the following?”* The activities included bathing, dressing, eating, getting in or out of chairs, walking, and using the toilet.³ Responses indicating either difficulty or inability to perform an ADL were categorized as “Impaired.”

The general health question asks: *“In general, would you say your health is....”* The comparative health question asks: *“In general, compared to other people your age, would you say your health is....”* For the general health and comparative health questions, the responses were dichotomized into the following categories: 1) “Excellent,” “Very good,” or “Good” and 2) “Fair” or “Poor.” The health transition question asks: *“Compared to one year ago, how would you rate your health in general now?”* For the health transition question, the responses were dichotomized into the categories of 1) “Much better now,” “Somewhat better now,” or “About the same now” and 2) “Somewhat worse now” or “Much worse now.”

A beneficiary was considered to have a positive depression screen if he or she answered “Yes” to any of the following three depression questions:

1. *“In the past year, have you had 2 weeks or more during which you felt sad, blue or depressed; or when you lost interest or pleasure in things that you usually cared about or enjoyed?”*
2. *“In the past year, have you felt depressed or sad much of the time?”*
3. *“Have you ever had 2 years or more in your life when you felt depressed or sad most days, even if you felt okay sometimes?”*

The number of chronic medical conditions was added together for 13 conditions and ranged from 0 to 13. The question prefacing each of the chronic medical conditions is: *“Has a doctor ever told you that you had...”*. The chronic medical conditions include: hypertension; angina pectoris or coronary artery disease; congestive heart failure (CHF); myocardial infarction or heart attack; other heart conditions; stroke; emphysema, asthma, and/or chronic obstructive pulmonary disease (COPD); inflammatory bowel disease, including Crohn’s disease and/or ulcerative

³ Possible responses for the ADL items included: “I am unable to do this activity,” “Yes, I have difficulty,” and “No, I do not have difficulty.”

colitis; arthritis of the hip or knee; arthritis of the hand or wrist; sciatica; diabetes, hyperglycemia, or glycosuria; and any cancer (other than skin cancer).

Age was categorized into the following six groups: 65-69; 70-74; 75-79; 80-84; 85-89; and 90 or more years. Race and marital status were dichotomized into two groups each: White and Nonwhite, and married and not married. Education was divided into three categories: some high school or less, high school graduate/GED,⁴ and some college/college graduate or more. Annual household income was also divided into three categories: income less than \$10,000, income between \$10,000 and \$19,999, and income of \$20,000 or more. Medicaid status had two categories: in Medicaid and not in Medicaid.

DATA ANALYSIS

The analysis was limited to seniors age 65 or older who completed a survey and responded to all six items of the ADL question. Proxy responses were included in the samples. For the purpose of this analysis, a completed survey was defined as one that could be used to estimate PCS and MCS scores using the 1998 norm-based standard scoring algorithm (NCQA, 2003).

Student's *t* tests were used to test for differences in the mean PCS and MCS scores, age, number of ADLs, and number of chronic medical conditions between the PACE and managed care groups. Chi-square tests of proportions were used to test for differences in all categorical variables between the groups. Analysis of variance (ANOVA) was used to test for differences between the subgroup means using multiple comparisons testing.

Traditional statistical measures produce numerous significant *p* values when large samples are compared, as is the case here. The question becomes, which of these statistically significant differences are large enough to be used in making policy decisions? Effect sizes were calculated for all means and proportions as an additional measure of the magnitude of the differences between the groups. The effect size for means is the standardized difference between two groups, and can be calculated by dividing the difference between two group means by a pooled standard deviation (Rosenthal & Rosnow, 1991). The effect size refers to the strength of the relationship in the population, and unlike a significance test, is relatively independent of the sample size. The most common standard for deciding which effect sizes are meaningful are Cohen's (1988) definition of small, medium, and large effect sizes, which we have adopted here. A small effect size is defined as greater than, or equal to, 0.20, but less than 0.50. A medium effect size is greater than, or equal to, 0.50, but less than 0.80. And a large effect size is greater than, or equal to, 0.80.

In addition to an overall comparison of PACE beneficiaries (PACE sample) and managed care beneficiaries (MA sample), three analyses were performed using matched samples from the MA sample to determine the independent and combined effects of matching MA beneficiaries to the PACE beneficiaries on age and ADL limitations. All MA matched samples were drawn using

⁴ Diploma for General Education

stratified random sampling at a ratio of 1:1 to the PACE sample. The first analysis matched the MA sample by age group (MA Age Group matched sample) to the PACE sample. The second analysis matched the MA sample by the type and number of ADL impairments (MA ADL matched sample) to the PACE sample. The third analysis matched the MA sample by both age group and the type and number of ADL impairments (MA Age & ADL matched sample) to the PACE sample.

All analyses were independently performed for *Cohorts II, III, IV, and V Baseline*. Since results across the cohorts were similar, the discussion will focus on *Cohort V Baseline*, which was the most recently administered baseline survey for which data were available. There were 96,807 MA respondents from a total of 154,016 eligible seniors (62.9 percent response rate) in 177 managed care plans, and 1,248 PACE respondents from a total of 3,926 eligible seniors (31.8 percent response rate) from 20 PACE plans that participated in *Cohort V Baseline* and met the selection criteria. Since four of the PACE respondents had a combined age group and ADL impairment pattern that could not be matched in the MA sample, these four were excluded from all analyses (adjusted PACE N = 1,244). The alpha level for all tests was 0.01. SAS[®] was used for all statistical analyses.⁵

⁵ SAS[®] is a registered trademark of the SAS Institute, Inc.

3

RESULTS

PACE SAMPLE VS. MA SAMPLE

There were significant differences in the characteristics of the *Cohort V Baseline* PACE and MA samples for all demographic factors shown in Table 1 ($p < 0.001$ for all analyses). The PACE sample was significantly older with a mean age of 81.1, compared to a mean age of 75.3 for the MA sample. There was a large effect size (0.9) for the difference in the overall mean ages. Thirty-three percent of the PACE sample and 9 percent of the MA sample were 85 years or older. Effect sizes for the differences within age groups between the samples were small in most categories. Respondents in the PACE sample were predominantly female (76.4 percent), Nonwhite (51.6 percent), widowed or not married (79.4 percent), had less than a high school degree (64.6 percent), had less than a \$10,000 annual household income (56.0 percent), and were enrolled in Medicaid (69.2 percent). Nearly two-thirds (59.1 percent) of the MA sample were female, 88.0 percent were White, 56.4 percent were married, 70.2 percent had a high school degree or college education, 85.2 percent had an annual household income of \$10,000 or more, and 95.6 percent were not enrolled in Medicaid. There were large effect sizes for the categorical differences between the samples for race, marital status, and Medicaid status. There were also large effect sizes for differences in income between samples for two of the three categories (income less than \$10,000 or income greater than, or equal to, \$20,000).

The *Cohort V Baseline* PACE and MA samples were also significantly different for all health status indicators shown in Table 2 ($p < 0.001$ for all analyses). PACE respondents were significantly more impaired in functional status than the MA respondents. The PACE sample had a mean of 3.9 ADL impairments compared to the MA sample mean of 1.0 ADL impairment. When assessing individual ADL impairments, the percentage of the PACE sample impaired on each ADL was significantly higher than for the MA sample. The most frequent ADL impairment for both samples was impairment with walking (83.8 percent for PACE compared to 35.6 percent for the MAs). There were large effect sizes for the differences in individual ADL impairment categories (0.8 to 1.3) and in the mean number of ADL impairments (1.8) between the samples.

The PACE sample had a mean of 4.1 chronic medical conditions, which was significantly greater than the MA sample mean of 2.7 conditions. The PACE sample responded “Fair” or “Poor” more frequently to the general health question (72.8 percent vs. 29.9 percent in the MA sample) and to the comparative health question (67.1 percent vs. 24.3 percent in the MA sample). PACE respondents were significantly more likely to respond “*Somewhat worse now*” or “*Much worse now*” to the health transition question (43.4 percent vs. 21.0 percent in the MA sample). In the PACE sample, 62.3 percent had a positive depression screen compared to 27.4 percent of the MA sample. A proxy completed the survey more frequently for beneficiaries of the PACE sample (75.5 percent vs. 12.1 percent in the MA sample). Effect sizes were large for categorical differences in the general health question, health comparison question, positive depression

screen, and proxy status. There were medium effect sizes for the differences in the health transition question and for the mean number of chronic medical conditions.

Tables 3 and 4 display the mean PCS and MCS scores overall, stratified by age group and by type and number of ADL impairments for the PACE sample and the MA sample. In Table 3, the overall mean PCS score was significantly lower for the PACE sample than for the MA sample (30.4 and 42.1, respectively). The multiple comparisons test for subgroup mean PCS differences indicated that the mean PCS scores were significantly lower for the PACE sample compared to the MA sample within all age group categories (overall $F(11, 98039) = 662.05, p < 0.001$). The mean PCS scores were significantly lower for the PACE sample than for the MA sample for all six types of individual ADL categories, as determined by t test results. The mean PCS scores for both samples were highest for those respondents who had no ADL impairments (42.1 for the PACE sample compared to 48.6 for the MA sample). The multiple comparisons test for subgroup mean PCS differences within the categories for the number of ADL impairments indicated that the mean PCS scores were significantly different in the PACE sample than in the MA sample in three categories: for those with zero, four, or six impairments (overall $F(13, 98037) = 7260.84, p < 0.001$).

In Table 4, the overall mean MCS score was significantly lower for the PACE sample than for the MA sample (42.2 and 51.6, respectively). The multiple comparisons test for subgroup mean MCS differences within the age groups indicated that the mean MCS scores were significantly lower for the PACE sample compared to the MA sample within all age group categories (overall $F(11, 98039) = 203.89, p < 0.001$). Mean MCS scores were also significantly lower for the PACE sample than for the MA sample within all age groups. Compared to the MA sample, beneficiaries of the PACE sample with ADL impairments had significantly lower mean MCS scores within all six individual ADL impairment categories as determined by t test results. The mean MCS scores for both samples were highest for those respondents who had no ADL impairments (49.3 for PACE compared to 54.3 for MA). The multiple comparisons test for the subgroup mean MCS differences within the categories for the number of ADL impairments indicated no differences in the mean MCS scores between the samples (overall $F(13, 98037) = 1373.43, p > 0.001$).

Effect sizes were large (1.0 and 0.9 respectively) for the differences in the overall mean PCS and MCS scores between the PACE and MA samples, as displayed in Tables 3 and 4. For differences between samples within the age group categories, effect sizes were medium to large. There were small effect sizes for the differences between samples in the mean PCS and MCS scores within most of the individual ADL impairment categories, and within many of the categories for the number of ADL impairments.

PACE SAMPLE VS. MA AGE & ADL MATCHED SAMPLE

Tables 5 and 6 show the mean PCS and MCS scores overall, stratified by age group and by type and number of ADL impairments for the PACE sample and the MA Age & ADL matched sample. As displayed in Table 5, the overall mean PCS score was not significantly different

between the samples (30.4 and 31.2, respectively). The interaction term for the multiple comparisons test for subgroup mean PCS score differences within the age group categories was not significant, $p = 0.271$, indicating there were no significant differences between the samples (overall $F(11, 2476) = 3.35, p < 0.001$). The mean PCS scores were not significantly different within any of the six individual ADL categories, as determined by t test results. The mean PCS scores for both samples were highest for those respondents who had no ADL impairments (42.1 for the PACE sample compared to 48.4 for the MA Age & ADL matched sample). The multiple comparisons test for subgroup mean PCS differences within the number of ADL impairment categories indicated that the mean PCS scores in the PACE sample were significantly different from the mean PCS scores in the MA Age & ADL matched sample in three categories: for those with zero, four, and six impairments (overall $F(13, 2474) = 89.62, p < 0.001$). Although the mean PCS scores for the PACE sample were lower for those with zero and six impairments, the mean PCS score was higher in the PACE sample for those with four ADL impairments when compared to the MA Age & ADL matched sample.

As shown in Table 6, the overall mean MCS score was significantly lower for the PACE sample than for the MA Age & ADL matched sample (42.2 and 44.5, respectively). The interaction term for the multiple comparisons test for subgroup mean MCS differences within the age group categories was not significant, $p = 0.904$, indicating there were no significant subgroup mean differences between the samples (overall $F(11, 2476) = 2.00, p = 0.025$). Compared to the MA Age & ADL matched sample, beneficiaries of the PACE sample with ADL impairments had significantly lower mean MCS scores for all six individual ADL impairments as determined by t test results. The mean MCS scores for both samples were highest for those respondents who had no ADL impairments (49.3 for the PACE sample compared to 52.2 for the MA Age & ADL matched sample). The interaction term for the multiple comparisons test for subgroup mean MCS differences within the number of ADL impairment categories was not significant, $p = 0.595$, indicating there were no subgroup mean differences between the samples (overall $F(13, 2474) = 27.03, p < 0.001$).

For the difference in the overall mean PCS score between the PACE and MA Age & ADL matched sample, no effect size was detected. The difference in the overall mean MCS score, however, had a small effect size (0.2), as displayed in Tables 5 and 6. There were small effect sizes for only two or three of the age group differences for either PCS or MCS. There were small effect sizes for the differences in the mean MCS scores within most of the categories for individual type and number of ADL impairments; however, for the differences in the mean PCS scores, no effect sizes were detected in many of these categories.

RESULTS OF COHORT V BASELINE SF-36 SUMMARY SCORES

Figure 1 displays the mean PCS and MCS scores for all samples in *Cohort V Baseline*: the PACE sample, the MA sample, the MA Age Group matched sample, the MA ADL matched sample, and the MA Age & ADL matched sample. As previously described, the overall mean PCS and MCS scores for the PACE sample were significantly lower than the corresponding mean scores in the MA sample. The overall mean PCS and MCS scores were also significantly lower for the

PACE sample than for the MA Age Group matched sample (39.2 and 50.6, respectively). When the MA sample was matched to the PACE sample by the type and number of ADL impairments, either with or without matching age, the overall mean PCS scores were not significantly lower for the PACE sample compared to the MA ADL matched sample (31.1) or the MA Age & ADL matched sample (31.2). The overall mean MCS scores were significantly lower for PACE compared to the MA ADL matched sample (44.3) or the MA Age & ADL matched sample (44.5).

RESULTS OF SF-36 SUMMARY SCORES FOR ALL COHORTS

Table 7 displays the mean PCS and MCS scores, the standard deviations, and the sample sizes for each PACE sample, each corresponding cohort MA sample, and all MA matched samples within *Cohorts II, III, IV, and V Baseline*. The mean PCS and MCS scores were relatively uniform within each sample across the four years of data collection. For PCS means, the largest spread of 0.8 occurred within the MA sample between *Cohorts II* and *V*. For MCS means, the largest spread of 0.6 occurred within the MA Age & ADL matched sample between *Cohorts II* and *V*.

4

DISCUSSION

Although there are some differences, the overall pattern of characteristics for the PACE and MA plan beneficiaries from the Medicare HOS in this analysis was similar to what has been reported from other sources. According to the National PACE Association (NPA), the typical PACE enrollee is an 81 year-old widow, living alone or with relatives, has several chronic medical conditions, may suffer some degree of cognitive impairment, and requires help with personal care and ADLs (NPA, 2002). Overall, most (53 percent) are from minority populations (Shannon & van Reenen, 1998). PACE enrollees receive help to perform many of the six major ADLs: walking (54 percent), getting in or out of chairs (48 percent), bathing (80 percent), dressing (69 percent), toileting (53 percent), and eating (36 percent); and have an average number of 3.4 ADL dependencies. PACE enrollees have been diagnosed with an average of 7.8 medical conditions (NPA, 2002). The HOS PACE beneficiaries tended to have a larger percentage with impairments of walking (83.8) and getting in or out of chairs (73.4), and they had a lower mean number (4.1) of chronic medical conditions compared to the NPA data.

In a comparison of Medicare Health Maintenance Organization (HMO) beneficiaries and PACE beneficiaries, Temkin-Greener, Meiners, and Gruenberg (2001) noted that PACE beneficiaries were older (34 percent in PACE were 85 or older compared to 6.8 percent in HMOs), more likely to be dual-eligible (95 percent in PACE vs. 4.4 percent in HMOs), experienced multiple limitations in ADLs (average of three in PACE vs. less than one in HMOs) and experienced a considerable degree of dementia (46 percent in PACE vs. 0.9 percent in HMOs).

A limitation of this study is the relatively small size of the PACE samples compared to the large MA samples within each of the cohorts. The method of 1:1 stratified random matching of the MA respondents to the PACE was selected to minimize this problem. In addition, effect sizes were utilized to give a more meaningful summary of the importance of the differences found.

Both the PACE and MA samples had high frequencies of non-respondents. In comparison to the PACE respondents in *Cohort V*, the mean age for PACE nonrespondents was slightly older (81.7); they were less often female (73.9 percent), more frequently Nonwhite (53.6 percent), and not married (81.9 percent), with less than a high school degree (67.8 percent). The characteristics of the MA nonrespondents compared to the MA respondents were similar. Plans with low response rates could have artificially higher scores because of greater non-response of those with lower functional status (Bierman, Lawrence, Haffer, & Clancy, 2001). Slightly higher percentages of high-risk beneficiaries in both nonrespondent groups indicated a potential source of bias relative to the somewhat healthier respondent groups in these samples.

The PACE sample had a high number of proxy-completed surveys. Elderly respondents are more likely to require assistance to respond or have cognitive impairments that can affect validity of responses (Bierman et al., 2001). There is evidence of different responses obtained from proxies

as opposed to self-reported responses. Neumann, Araki, and Gutterman (2000) suggest that more observable conditions tend to be reported more accurately with proxy-completed responses, while less observable conditions such as those involving emotional states tend to be underestimated. Other research suggests that proxy-completed responses may underestimate the prevalence of certain health conditions (Shields, 2000). Future research utilizing the PACE program data is indicated to determine the extent of the bias with proxy completion of surveys.

5

CONCLUSIONS

In summary, the MA respondents in HOS *Cohorts II, III, IV, and V Baseline* were younger, had fewer ADL impairments, and had better functional status, as measured by higher PCS and MCS scores, compared to the PACE respondents in the corresponding cohort. Similar results were found when the MA respondents were randomly matched by age group to the PACE respondents.

Within each cohort, when MA respondents were matched to PACE respondents by the type and number of ADL impairments, either with or without matching by age group, the mean PCS scores for the MA matched samples were not significantly higher compared to the corresponding PACE samples. Although the mean MCS scores were significantly higher for the MA matched samples compared to the PACE sample within *Cohort V*, the small effect size indicated a weak relationship. The results for *Cohorts II, III, and IV* indicated there were no significant differences in MCS scores between these samples and the corresponding PACE samples. These results suggest that comparisons between beneficiaries of PACE and MA plans should be controlled for both age and functional limitations.

6

REFERENCES

- Bierman, A. S., Lawrence, W. F., Haffer, S. C., & Clancy, C. M. (2001). "Functional Health Outcomes as a Measure of Health Care Quality for Medicare Beneficiaries." *Health Services Research*, 35(6), 90-109.
- Bodenheimer, T. (1999). Long-Term Care For Frail Elderly People–The On Lok Model. *The New England Journal of Medicine*, 34(17), 1324-1328.
- Centers for Medicare & Medicaid Services. (2001). Protecting Medicare Beneficiaries When Their Medicare + Choice Organization Withdraws. September 1, 2001. Retrieved on October 7, 2005, from <http://www.cms.hhs.gov/media/press/release.asp?Counter=402>.
- Centers for Medicare & Medicaid Services. (2002). Program of All-Inclusive Care for the Elderly (PACE) History. August 6, 2002. Retrieved on October 7, 2005, from <http://cms.hhs.gov/pace/pacegen.asp>.
- Centers for Medicare & Medicaid Services. (2005). Alternatives to Nursing Home Care: About PACE. Retrieved on October 7, 2005, from <http://www.medicare.gov/Nursing/Alternatives/Pace.asp>.
- Chatterji, P., Burstein, N. R., Kidder, D., & White, A. J. (1998). Evaluation of the Program of All-Inclusive Care for the Elderly (PACE) Demonstration: The Impact of PACE on Participant Outcomes. Cambridge, MA: Abt Associates.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Second Edition. Hillsdale, NJ: Erlbaum.
- Department of Health and Human Services. (1999). Health Care Financing Administration 42 CFR Parts 460, et al. *Medicare and Medicaid Programs; Program of All-Inclusive Care for the Elderly (PACE)*; Final Rule. *Federal Register*, 64 (226), 66233-66290, November 24, 1999.
- Department of Health and Human Services. (2003). Centers for Medicare & Medicaid Services 42 CFR Parts 409, et al. *Medicare Program; Modifications to Managed Care Rules*. *Federal Register*, 68(163), 50840-50859, August 22, 2003.
- Eng, C., Pedulla, J., Eleazer, G. P., McCann, R., & Fox, N. (1997). Program of All-inclusive Care for the Elderly (PACE): An Innovative Model of Integrated Geriatric Care and Financing. *Journal of the American Geriatrics Society*, 45, 223-232.

- McHorney, C. A., Ware, J. E., Lu, J. F. R., & Sherbourne, C. D. (1994). The MOS 36-Item Short Form Health Survey (SF-36): III Test of data quality, scaling assumptions and reliability across diverse patient groups. *Medical Care*, 32, 40-66.
- National Committee for Quality Assurance. (2003). *HEDIS® 2003, Volume 6: Specifications for the Medicare Health Outcomes Survey*. Washington DC: National Committee for Quality Assurance Publication.
- National PACE Association. (2002). PACE Profile: PACE Enrollee Characteristics. Retrieved on October 7, 2005 from <http://www.npaonline.org/content/research/profile/profile3.asp>.
- Neumann, P. J., Araki, S. S., & Gutterman, E. M. (2000). The Use of Proxy-completed Responses in Studies of Older Adults: Lessons, Challenges, and Opportunities. *Journal of the American Geriatrics Society*, 48, 1646-1654.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of Behavioral Research: Methods and Data Analysis*. Second Edition. McGraw-Hill, Inc.
- Shannon, K. & van Reenen, C. (1998). Catholic Health Association of the United States: Health Progress. Retrieved on October 7, 2005 from <http://www.chausa.org/PUBS/PUBSART.ASP?ISSUE=HP9809&ARTICLE=C>.
- Shields, M. (2000). Proxy Reporting in the National Population Health Survey. *Health Reports*, 12, 21-39.
- Temkin-Greener, H., Meiners, M. R., & Gruenberg, L. (2001). PACE and the Medicare + Choice Risk-Adjusted Payment Model. *Inquiry*, 38, 60-72.
- Weiland, D., Lamb, V. L. Sutton, S. R., Boland, R., Clark, M., Friedman, S., Brummel-Smith, K., & Eleazer, G. P. (2000). Hospitalization in the Program of All-Inclusive Care for the Elderly (PACE): Rates, Concomitants, and Predictors. *Journal of the American Geriatrics Society*, 48(11), 1373-1380.

APPENDIX

Table 1: Cohort V Baseline Demographic Summary for the PACE Sample and the MA Sample						
Sample ¹	PACE (N=1,244)		MA (N=96,807)		P Value ²	Effect Size ³
Mean Age (SD)	81.1	(7.9)	75.3	(6.5)	<0.001	0.9
Percentage by Age Group (N)					<0.001	
Age 65 - 69	8.4	(105)	24.9	(24,136)		0.5
Age 70 - 74	16.6	(207)	28.7	(27,784)		0.3
Age 75 - 79	20.2	(251)	22.9	(22,116)		0.1
Age 80 - 84	21.5	(267)	14.5	(14,050)		0.2
Age 85 - 89	19.5	(242)	6.5	(6,330)		0.4
Age 90 or more	13.8	(172)	2.5	(2,391)		0.4
Percentage by Gender (N)					<0.001	0.4
Male	23.6	(294)	40.9	(39,632)		
Female	76.4	(950)	59.1	(57,175)		
Percentage by Race (N)					<0.001	0.9
White	48.4	(602)	88.0	(85,221)		
Nonwhite	51.6	(642)	12.0	(11,586)		
Percentage by Marital Status (N)					<0.001	0.8
Married	20.6	(253)	56.4	(53,941)		
Not married	79.4	(975)	43.6	(41,723)		
Percentage by Education (N)					<0.001	
Some high school or less	64.6	(780)	29.8	(28,336)		0.7
High school graduate/GED ⁴	22.8	(275)	37.0	(35,262)		0.3
Some college/College graduate or more	12.6	(152)	33.2	(31,618)		0.5
Household Income (N)					<0.001	
Income < \$10,000	56.0	(525)	14.8	(11,333)		0.9
Income ≥ \$10,000 and ≤ \$19,999	27.0	(253)	30.5	(23,391)		0.1
Income ≥ \$20,000	17.0	(159)	54.7	(41,987)		0.8
Percentage by Medicaid Status (N)					<0.001	1.5
Medicaid	69.2	(861)	4.4	(4,218)		
No Medicaid	30.8	(383)	95.6	(92,589)		

¹ The PACE sample and the MA sample included respondents age 65 or older, with PCS and MCS scores, and responses to all six Activities of Daily Living (ADL) items.

² The *t* test was performed for the difference in mean age, and Chi-square tests of proportions were performed for all categorical variables.

³ Effect size calculation for means: $ES = |O_1 - O_2| / s$ pooled;
 Effect size calculation for proportions: $ES = |(2\arcsin \sqrt{P_1}) - (2\arcsin \sqrt{P_2})|$.

⁴ Diploma of General Education.

**Table 2: Cohort V Baseline Health Status Summary
 for the PACE Sample and the MA Sample**

Sample¹	PACE (N=1,244)	MA (N=96,807)	P Value²	Effect Size³
Percentage by Type of ADL Impairment (N)⁴				
Bathing	76.2 (948)	14.6 (14,102)	<0.001	1.3
Dressing	69.7 (867)	11.9 (11,548)	<0.001	1.3
Eating	34.4 (428)	5.8 (5,582)	<0.001	0.8
Getting in or out of chairs	73.4 (913)	27.3 (26,428)	<0.001	1.0
Walking	83.8 (1,043)	35.6 (34,430)	<0.001	1.0
Using the toilet	53.3 (663)	8.4 (8,120)	<0.001	1.0
Mean Number of ADL Impairments (SD)	3.9 (2.0)	1.0 (1.6)	<0.001	1.8
Percentage by Number of ADL Impairments (N)			<0.001	
Zero	8.8 (109)	58.0 (56,143)		1.1
One	8.0 (100)	14.8 (14,358)		0.2
Two	9.7 (120)	12.3 (11,860)		0.1
Three	9.7 (121)	4.9 (4,703)		0.2
Four	13.8 (171)	3.8 (3,649)		0.4
Five	21.1 (263)	3.2 (3,137)		0.6
Six	28.9 (360)	3.1 (2,957)		0.8
Mean Number of Chronic Medical Conditions (SD)	4.1 (2.4)	2.7 (2.0)	<0.001	0.7
Percentage by General Health (N)			<0.001	0.9
Excellent/Very good/Good	27.2 (336)	70.2 (67,408)		
Fair/Poor	72.8 (898)	29.9 (28,678)		
Percentage by Health Compared to One Year Ago (N)			<0.001	0.5
Much better/Somewhat better/About the same	56.6 (698)	79.0 (75,963)		
Somewhat worse/Much worse	43.4 (536)	21.0 (20,163)		
Percentage by Health Compared to Peers (N)			<0.001	0.9
Excellent/Very good/Good	32.9 (404)	75.7 (72,458)		
Fair/Poor	67.1 (824)	24.3 (23,210)		
Percentage by Positive Depression Screen (N)			<0.001	1.5
Positive	62.3 (760)	27.4 (26,078)		
Negative	37.7 (460)	72.6 (68,975)		
Percentage by Person Who Completed the Survey (N)			<0.001	1.4
Self	24.5 (277)	87.9 (78,568)		
Others	75.5 (855)	12.1 (10,842)		

¹ The PACE sample and the MA sample included respondents age 65 or older, with PCS and MCS scores, and responses to all six Activities of Daily Living (ADL) items.

² The *t* tests were performed for the differences in mean number of ADL impairments and mean number of chronic conditions, and Chi-square tests of proportions were performed for all categorical variables.

³ Effect size calculation for means: $ES = |O_1 - O_2| / s$ pooled;
 Effect size calculation for proportions: $ES = |(2\arcsin \sqrt{P_1}) - (2\arcsin \sqrt{P_2})|$.

⁴ Responses indicating either difficulty or inability to perform the ADL were categorized as impaired.

Table 3: Cohort V Baseline Physical Component Summary (PCS) Scores for the PACE Sample and the MA Sample

Sample ¹	PCS Scores			
	PACE (N=1,244)	MA (N=96,807)	P Value ⁵	Effect Size ⁶
Mean Scores By Total (SD)	30.4 (9.2)	42.1 (11.5)	<0.001	1.0
Mean Scores by Age Group (SD)²				
Age 65 - 69	30.1 (9.5)	45.0 (11.1)	<0.001	1.3
Age 70 - 74	31.2 (9.4)	43.6 (11.2)	<0.001	1.1
Age 75 - 79	32.2 (9.8)	41.3 (11.3)	<0.001	0.8
Age 80 - 84	30.8 (8.8)	39.0 (11.3)	<0.001	0.7
Age 85 - 89	29.1 (8.4)	36.9 (11.0)	<0.001	0.7
Age 90 or more	28.5 (9.0)	34.1 (10.4)	<0.001	0.5
Mean Scores by Type of ADL Impairment (SD)³				
Bathing	28.5 (8.2)	29.4 (8.5)	0.004	0.1
Dressing	28.2 (8.3)	29.1 (8.6)	0.004	0.1
Eating	27.5 (8.2)	31.7 (10.0)	<0.001	0.4
Getting in or out of chairs	28.1 (8.1)	31.8 (8.9)	<0.001	0.4
Walking	28.7 (8.2)	31.9 (8.5)	<0.001	0.4
Using the toilet	27.3 (7.9)	29.9 (9.2)	<0.001	0.3
Mean Scores by Number of ADL Impairments (SD)⁴				
Zero	42.1 (8.9)	48.6 (8.3)	<0.001	0.8
One	35.6 (8.2)	37.8 (8.6)	0.010	0.3
Two	33.5 (7.8)	32.8 (7.8)	0.333	0.1
Three	31.1 (7.4)	30.1 (7.6)	0.201	0.1
Four	29.5 (8.2)	27.8 (7.2)	0.008	0.2
Five	27.9 (8.2)	27.2 (7.3)	0.187	0.1
Six	26.5 (7.7)	29.9 (10.0)	<0.001	0.3

¹ The PACE sample and the MA sample included respondents age 65 or older, with PCS and MCS scores, and responses to all six Activities of Daily Living (ADL) items.

² ANOVA was performed to test subgroup mean PCS differences within the age groups, and indicated that the mean PCS scores were significantly different at the $p < 0.01$ level between the PACE sample and the MA sample within all age group categories: Overall $F(11, 98039) = 662.05, p < 0.001$; Group ($p < 0.001$); Age Group ($p < 0.001$); and Group by Age Group ($p < 0.001$).

³ Responses indicating either difficulty or inability to perform the ADL were categorized as impaired.

⁴ ANOVA was performed to test subgroup mean PCS differences within number of ADL impairments groups, and indicated that the mean PCS scores were significantly different at the $p < 0.01$ level between the PACE sample and the MA sample for zero, four, and six ADL impairments: Overall $F(13, 98037) = 7260.84, p < 0.001$; Group ($p < 0.001$); Number of ADL Impairments ($p < 0.01$); and Group by Number of ADL Impairments ($p < 0.001$).

⁵ The t tests were performed for the differences of sample means.

⁶ Effect size calculation for means: $ES = |O_1 - O_2| / s$ pooled;
 Effect size calculation for proportions: $ES = |(2\arcsin \sqrt{P_1}) - (2\arcsin \sqrt{P_2})|$.

**Table 4: Cohort V Baseline Mental Component Summary (MCS) Scores
 for the PACE Sample and the MA Sample**

Sample ¹	MCS Scores			
	PACE (N=1,244)	MA (N=96,807)	P Value ⁵	Effect Size ⁶
Mean Scores By Total (SD)	42.2 (13.0)	51.6 (10.4)	<0.001	0.9
Mean Scores by Age Group (SD)²				
Age 65 - 69	41.8 (12.8)	52.8 (9.7)	<0.001	1.1
Age 70 - 74	42.2 (12.7)	52.4 (10.1)	<0.001	1.0
Age 75 - 79	42.6 (12.9)	51.3 (10.6)	<0.001	0.8
Age 80 - 84	42.1 (12.5)	50.4 (11.0)	<0.001	0.8
Age 85 - 89	42.1 (13.8)	49.2 (11.3)	<0.001	0.6
Age 90 or more	42.5 (13.3)	48.4 (12.0)	<0.001	0.5
Mean Scores by Type of ADL Impairment (SD)³				
Bathing	40.4 (12.7)	43.4 (12.3)	<0.001	0.2
Dressing	40.3 (12.9)	43.0 (12.4)	<0.001	0.2
Eating	36.9 (12.2)	40.9 (12.3)	<0.001	0.3
Getting in or out of chairs	40.4 (12.9)	47.0 (12.2)	<0.001	0.5
Walking	41.3 (13.0)	47.7 (12.1)	<0.001	0.5
Using the toilet	39.1 (12.6)	42.5 (12.5)	<0.001	0.3
Mean Scores by Number of ADL Impairments (SD)⁴				
Zero	49.3 (11.8)	54.3 (8.2)	<0.001	0.6
One	48.8 (11.9)	51.4 (10.6)	0.007	0.2
Two	46.2 (11.7)	49.6 (11.3)	<0.001	0.3
Three	44.2 (12.3)	46.1 (11.8)	0.028	0.2
Four	42.4 (12.4)	44.2 (11.8)	0.017	0.2
Five	41.7 (13.1)	41.8 (12.0)	0.965	0.0
Six	36.6 (12.0)	39.3 (12.7)	<0.001	0.2

¹ The PACE sample and the MA sample included respondents age 65 or older, with PCS and MCS scores, and responses to all six Activities of Daily Living (ADL) items.

² ANOVA was performed to test subgroup mean MCS differences within the age groups, and indicated that the mean MCS scores were significantly different at the $p < 0.01$ level between the PACE sample and the MA sample within all age group categories: Overall $F(11, 98039) = 203.89, p < 0.001$; Group ($p < 0.001$); Age Group ($p < 0.001$); and Group by Age Group ($p < 0.001$).

³ Responses indicating either difficulty or inability to perform the ADL were categorized as impaired.

⁴ ANOVA was performed to test subgroup mean MCS differences within number of ADL impairments groups, and indicated that the mean MCS scores were significantly different at the $p < 0.01$ level between the PACE sample and the MA sample for zero, one, two, and six ADL impairments: Overall $F(13, 98037) = 1373.43, p < 0.001$; Group ($p < 0.001$); Number of ADL Impairments ($p < 0.001$); and Group by Number of ADL Impairments ($p < 0.001$).

⁵ The t tests were performed for the difference of sample means.

⁶ Effect size calculation for means: $ES = |O_1 - O_2| / s \text{ pooled}$;

Effect size calculation for proportions: $ES = |(2\arcsin \sqrt{P_1}) - (2\arcsin \sqrt{P_2})|$.

**Table 5: Cohort V Baseline Physical Component Summary (PCS) Scores
 for the PACE Sample and the MA Age & ADL Matched Sample**

Sample ¹	PCS Scores			
	PACE (N=1,244)	MA Age & ADL Matched ² (N=1,244)	P Value ⁵	Effect Size ⁶
Mean Scores By Total (SD)	30.4 (9.2)	31.2 (10.1)	0.056	0.1
Mean Scores by Age Group (SD)²				
Age 65 - 69	30.1 (9.5)	32.1 (10.7)	0.128	0.2
Age 70 - 74	31.2 (9.4)	32.5 (11.4)	0.144	0.1
Age 75 - 79	32.2 (9.8)	31.2 (9.5)	0.256	0.1
Age 80 - 84	30.8 (8.8)	31.3 (9.9)	0.585	0.1
Age 85 - 89	29.1 (8.4)	30.6 (10.0)	0.078	0.2
Age 90 or more	28.5 (9.0)	29.5 (8.9)	0.361	0.1
Mean Scores by Type of ADL Impairment (SD)³				
Bathing	28.5 (8.2)	28.5 (8.4)	0.968	0.0
Dressing	28.2 (8.3)	28.3 (8.5)	0.754	0.0
Eating	27.5 (8.2)	28.7 (9.3)	0.053	0.1
Getting in or out of chairs	28.1 (8.1)	28.4 (8.4)	0.491	0.0
Walking	28.7 (8.2)	28.8 (8.3)	0.760	0.0
Using the toilet	27.3 (7.9)	28.3 (8.6)	0.042	0.1
Mean Scores by Number of ADL Impairments (SD)⁴				
Zero	42.1 (8.9)	48.4 (7.6)	<0.001	0.8
One	35.6 (8.2)	38.3 (7.9)	0.019	0.3
Two	33.5 (7.8)	33.2 (7.3)	0.790	0.0
Three	31.1 (7.4)	30.1 (7.2)	0.360	0.1
Four	29.5 (8.2)	27.1 (6.8)	0.005	0.3
Five	27.9 (8.2)	27.0 (7.3)	0.203	0.1
Six	26.5 (7.7)	28.6 (9.4)	<0.001	0.2

¹ The PACE sample and the MA sample included respondents age 65 or older, with PCS and MCS scores, and responses to all six Activities of Daily Living (ADL) items.

² ANOVA was performed to test subgroup mean PCS differences within the age groups, and indicated that mean PCS scores were not significantly different at the $p < 0.01$ level between the PACE sample and MA Age & ADL matched sample, due to a non-significant interaction term: Overall $F(11, 2476) = 3.35, p < 0.001$; Group ($p = 0.028$); Age Group ($p < 0.001$); and Group by Age Group ($p = 0.271$).

³ Responses indicating either difficulty or inability to perform the ADL were categorized as impaired.

⁴ ANOVA was performed to test subgroup mean PCS differences within ADL impairment groups, and indicated that mean PCS scores were significantly different at the $p < 0.01$ level between the PACE sample and the MA Age & ADL matched sample for zero, four, and six ADL impairments: Overall $F(13, 2474) = 89.62, p < 0.001$; Group ($p = 0.009$); Number of ADL Impairments ($p < 0.001$); and Group by Number of ADL Impairments ($p < 0.001$).

⁵ The t tests were performed for the difference of sample means.

⁶ Effect size calculation for means: $ES = |O_1 - O_2| / s \text{ pooled}$;
 Effect size calculation for proportions: $ES = |(2\arcsin \sqrt{P_1}) - (2\arcsin \sqrt{P_2})|$.

**Table 6: Cohort V Baseline Mental Component Summary (MCS) Scores
 for the PACE Sample and the MA Age & ADL Matched Sample**

Sample ¹	MCS Scores			
	PACE (N=1,244)	MA Age & ADL Matched (N=1,244)	P Value ⁵	Effect Size ⁶
Mean Scores By Total (SD)	42.2 (13.0)	44.5 (12.3)	<0.001	0.2
Mean Scores by Age Group (SD)²				
Age 65 - 69	41.8 (12.8)	44.9 (11.6)	0.073	0.3
Age 70 - 74	42.2 (12.7)	45.1 (12.7)	0.019	0.2
Age 75 - 79	42.6 (12.9)	44.2 (12.4)	0.155	0.1
Age 80 - 84	42.1 (12.5)	44.9 (12.6)	0.009	0.2
Age 85 - 89	42.1 (13.8)	44.0 (12.2)	0.099	0.1
Age 90 or more	42.5 (13.3)	43.9 (12.3)	0.278	0.1
Mean Scores by Type of ADL Impairment (SD)³				
Bathing	40.4 (12.7)	42.6 (12.2)	<0.001	0.2
Dressing	40.3 (12.9)	42.1 (12.3)	0.003	0.1
Eating	36.9 (12.2)	39.8 (12.2)	0.001	0.2
Getting in or out of chairs	40.4 (12.9)	42.8 (12.4)	<0.001	0.2
Walking	41.3 (13.0)	43.6 (12.5)	<0.001	0.2
Using the toilet	39.1 (12.6)	41.1 (12.2)	0.003	0.2
Mean Scores by Number of ADL Impairments (SD)⁴				
Zero	49.3 (11.8)	52.2 (8.6)	0.072	0.3
One	48.8 (11.9)	50.9 (11.0)	0.197	0.2
Two	46.2 (11.7)	48.8 (11.4)	0.089	0.2
Three	44.2 (12.3)	46.0 (11.2)	0.248	0.2
Four	42.4 (12.4)	45.7 (11.9)	0.009	0.3
Five	41.7 (13.1)	42.2 (11.5)	0.656	0.0
Six	36.6 (12.0)	39.6 (12.5)	<0.001	0.2

¹ The PACE sample and the MA sample included respondents age 65 or older, with PCS and MCS scores, and responses to all six Activities of Daily Living (ADL) items.

² ANOVA was performed to test subgroup mean MCS differences within the age groups, and indicated that mean MCS scores were not significantly different at the $p < 0.01$ level between the PACE sample and the MA Age & ADL matched sample, due to a non-significant overall F test result: Overall $F(11, 2476) = 2.00, p = 0.025$; Group ($p < 0.001$); Age Group ($p = 0.990$); and Group by Age Group ($p = 0.904$).

³ Responses indicating either difficulty or inability to perform the ADL were categorized as impaired.

⁴ ANOVA tests were performed to test subgroup mean MCS differences within number of ADL impairments groups, and indicated that mean MCS scores were not significantly different at the $p < 0.01$ level between the PACE sample and the MA Age & ADL matched sample, due to a non-significant interaction term, Overall $F(13, 2474) = 27.03, p < 0.001$; Group ($p < 0.001$); Number of ADL Impairments ($p < 0.001$); and Group by Number of ADL Impairments ($p = 0.595$).

⁵ The t tests were performed for the difference of sample means.

⁶ Effect size calculation for means: $ES = |O_1 - O_2| / s_{\text{pooled}}$;
 Effect size calculation for proportions: $ES = |(2\arcsin \sqrt{P_1}) - (2\arcsin \sqrt{P_2})|$.

Table 7: Cohorts II, III, IV, and V Baseline SF-36 Summary Scores								
	Cohort II		Cohort III		Cohort IV		Cohort V	
	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
PCS by Sample¹								
PACE Sample	31.0 (9.5)	1,453	30.8 (9.4)	1,428	30.7 (9.2)	1,126	30.4 (9.2)	1,244
MA Sample	42.9 (11.4)	171,813	42.5 (11.5)	181,407	42.3 (11.5)	110,642	42.1 (11.5)	96,807
MA Age Group Matched Sample ²	39.8 (11.4)	1,453	39.4 (11.8)	1,428	39.1 (11.5)	1,126	39.2 (11.5)	1,244
MA ADL Matched Sample ³	31.6 (10.5)	1,453	31.6 (10.4)	1,428	31.8 (10.8)	1,126	31.1 (10.2)	1,244
MA Age & ADL Matched Sample ⁴	31.2 (10.3)	1,453	31.2 (10.1)	1,428	31.6 (10.0)	1,126	31.2 (10.1)	1,244
MCS by Sample								
PACE Sample	42.8 (12.9)	1,453	42.5 (13.2)	1,428	42.6 (12.8)	1,126	42.2 (13.0)	1,244
MA Sample	51.8 (10.3)	171,813	51.7 (10.5)	181,407	51.8 (10.4)	110,642	51.6 (10.4)	96,807
MA Age Group Matched Sample	50.5 (11.0)	1,453	50.1 (11.5)	1,428	50.2 (11.2)	1,126	50.6 (10.9)	1,244
MA ADL Matched Sample	43.6 (12.7)	1,453	43.7 (12.8)	1,428	44.3 (12.6)	1,126	44.3 (12.6)	1,244
MA Age & ADL Matched Sample	43.9 (12.7)	1,453	43.2 (13.0)	1,428	44.3 (12.9)	1,126	44.5 (12.3)	1,244

¹ All samples included respondents age 65 or older with PCS and MCS scores and responses to all six Activities of Daily Living (ADL) items.

² MA respondents who were matched to PACE respondents by age group.

³ MA respondents who were matched to PACE respondents by type and number of ADL impairments.

⁴ MA respondents who were matched to PACE respondents by age group, and by type and number of ADL impairments.