LONG-TERM FUNCTIONING AND WELL-BEING OUTCOMES ASSOCIATED WITH PHYSICAL ACTIVITY AND EXERCISE IN PATIENTS WITH CHRONIC CONDITIONS IN THE MEDICAL OUTCOMES STUDY

ANITA L. STEWART,* RON D. HAYS,* KENNETH B. WELLS,2,3 WILLIAM H. ROGERS,2 KAREN L. SPRITZER2 and SHELDON GREENFIELD4

1University of California San Francisco, Institute for Health & Aging, San Francisco, CA 94143, 2RAND, Santa Monica, CA, 3UCLA Neuropsychiatric Institute, UCLA School of Medicine, Los Angeles, CA and 4The Health Institute, New England Medical Center, Boston, MA, U.S.A.

(Received 1 February 1994)

Abstract—This study was carried out to determine whether levels of physical activity of patients with various chronic diseases are associated with subsequent functioning and well-being. It was an observational 2-year longitudinal design. The setting was offices of medical and mental health practices within health maintenance organizations, large multispecialty groups, and solo practices or small single-specialty group practices in three U.S. cities. Included in the study were 1758 adult patients with one or more of the following: diabetes, hypertension, congestive heart failure, recent myocardial infarction, depressive symptoms, or current depressive disorder. Outcome measures included physical, role, and social functioning; energy/fatigue; pain intensity; sleep problems; depressed affect, anxiety, positive affect, and overall psychological distress/well-being; health distress; and current health perceptions. Cross-sectional (baseline), 2-year endpoint, and change score relationships were evaluated between baseline levels of physical activity and each outcome, controlling for chronic conditions, comorbidity, smoking, alcohol use, overweight, self-reported adherence, and other patient and study characteristics. Higher baseline levels of exercise were uniquely associated with better functioning and well-being at baseline and 2 years later for some measures. The magnitude of the differences varied by disease group, but tended to be between 0.17 and 0.39 of the baseline SD. Greater levels of exercise are associated with feeling and functioning better for patients with chronic conditions over a 2-year period, suggesting that this is a fruitful area for further study using controlled interventions.

Exercise Physical activity Chronic disease Depression Functional status Health status Outcomes Quality of life Health-related quality of life

INTRODUCTION

Exercise and physical activity may prevent a variety of future health problems [1]. People who exercise regularly are more likely to live longer [2–5], less likely to develop heart disease [6, 7], and less likely to develop hypertension [6]. Various physiological benefits, such as improved cardiovascular functioning, have also been found in general populations [8].
Less is known about the benefits of physical activity for those who already have a chronic medical or mental health condition. Because of the substantial prevalence of chronic conditions [9] and their associated health costs due to use of services and morbidity, whether exercise reduces the morbidity associated with chronic conditions is of considerable policy interest. Further, this issue is of potential relevance to clinicians.

Some benefits of physical activity have been observed for individuals with chronic illnesses in terms of various physiological and other clinical indicators of disease status [10–18]. However, the primary goal of treatment of patients with chronic medical conditions is to enhance functioning and well-being [19,20]. To the extent that exercise improves functioning and well-being, increasing physical activity may be one means that clinicians can use to achieve this primary treatment goal.

There are important limitations to existing studies of this issue. First, most studies focus on people with only one medical disease or condition, but many who have one condition also have other medical conditions [21,22] or a medical condition and depression [23]. Thus, it is important to evaluate the benefits of exercise for people with multiple chronic conditions. Second, previous studies tend to examine only a few outcomes, typically purely psychological domains such as emotional well-being. For patients with chronic medical conditions, however, benefits of exercise may be more evidenced in domains such as physical, role, and social functioning or pain and fatigue, because the limitations associated with medical conditions are greatest for these domains [22]. Third, most studies focus on the benefits of aerobic or endurance training, a level of activity which may be too strenuous for persons with severe chronic conditions. Thus, some attention needs to be given to less strenuous forms of exercise such as walking.

In this paper, we study patients with a spectrum of diseases and disease severity, including depression; we examine a variety of domains of functioning and well-being; and we examine associations of functioning and well-being with several types of and levels of physical activity, including walking.

METHODS

Sample and study design

The data are from patients enrolled in the longitudinal portion of the Medical Outcomes Study (MOS), a large national study of the processes and outcomes of medical care in three systems of care: health maintenance organizations, large multispecialty groups, and solo or small single-specialty fee-for-service practice. Patients had one or more chronic medical conditions (hypertension, diabetes, a recent myocardial infarction, congestive heart failure) and/or depression. Details regarding the study are provided elsewhere [24–26].

To summarize briefly, three study sites (Boston, Chicago, Los Angeles) with mature forms of each system of care were chosen from Standard Metropolitan Statistical Areas. Within each system, a representative sample of physicians (general internists, family physicians, cardiologists, endocrinologists, diabetologists, psychiatrists), psychologists, other mental health providers, and nurse practitioners was selected (N = 523 providers). Among participating providers, a representative cross-section of their patients was screened for possible inclusion in the MOS panel. All adult patients visiting one of the MOS clinicians during an average 9-day period were screened. Patients who were under age 18, did not speak English, or were physically impaired in a way that would prohibit completing forms (e.g. blind) were excluded.

Patients with hypertension were eligible for the longitudinal study based on systolic and diastolic blood pressure readings reported by physicians. Patients with diabetes mellitus were identified through physician reports of the presence of diabetes; type of diabetes was determined through information on age of onset and insulin use—the latter was determined from patient report. Heart disease patients were identified through physician report of having suffered a myocardial infarction within the prior 12 months or having congestive heart failure. Patients with a level of depressive symptoms indicating a high risk for affective disorder were identified using an 8-item screening questionnaire [27]. Those exceeding an established cut-point had about a 35% chance of having current major depression or dysthymia (i.e. positive predictive value). Patients who enrolled in the panel were better educated than those who did not enroll [24].
Our analyses included individuals who enrolled in the longitudinal study, completed baseline questionnaires containing measures of physical activity and functional status, and also completed a 2-year follow-up questionnaire (N = 1758).

**Measures of physical activity and exercise**

The MOS defined physical activity as “bodily movement produced by skeletal muscles that results in energy expenditure” [28] and exercise as physical activity engaged in during leisure time. Physical activity was assessed by 11 items in the baseline questionnaire. This set of items was developed specifically for the MOS, but was based in part on prior measures [29]. Eight measures were derived from the 11 items: frequency and total time spent exercising, frequency and total time spent walking, amount of energy expended in leisure activities (achieved by weighting the amount of time spent in various leisure activities by the metabolic cost of the activities), amount of energy expended in work activities, walking pace, and perceived level of physical activity [29]. For this study, we selected a subset of three of the eight measures that were the most conceptually distinct and that were not highly correlated: total time spent exercising, total time spent walking, and perceived level of physical activity [29]. For this study, we selected a subset of three of the eight measures that were the most conceptually distinct and that were not highly correlated: total time spent exercising, total time spent walking, and perceived level of physical activity. Correlations among these three measures ranged from 0.25 to 0.44. Correlations among all eight measures ranged from 0.00 to 0.71 (median 0.31). All analyses in this paper focus on the subset of three physical activity measures.

**Measures of other health behaviors**

Smoking status is assessed by self-report in terms of whether the person is a current or past smoker or has never smoked. Overweight is assessed in terms of self-reported weight, adjusted for height (wt/ht2 for men, wt/ht1.5 for women) according to standard formulas [30]. A cutpoint on the indicator identifies those who are either moderately or severely overweight [30]. Alcohol use is assessed in terms of an indicator of problem drinking [31]. Patients were classified as current problem drinkers if they reported that they had a drink of beer, wine or hard liquor within the last 3 months and in addition, at least one of five symptoms or negative consequences of drinking (e.g. had five or more drinks at a time at a single sitting). Generic or general measures of adherence assess attitudes toward adherence regardless of the type of treatment recommended [32]. The MOS general adherence scales used five items administered with six responses ranging from none of the time to all of the time [33]. The internal-consistency reliability of the scale was acceptable (α = 0.81).

**Measures of functioning and well-being**

Functioning refers to people’s ability to perform their daily tasks and activities—behaviors that are usually observable. Well-being refers to more subjective internal states including how people feel physically and emotionally and how they think and feel about their health. We assessed 13 measures of functioning and well-being: physical functioning, role limitations (due to physical health and due to emotional problems), social activities limitations due to health, energy/fatigue, pain severity, sleep problems, health distress, current health perceptions, and psychological distress/well-being (depression-behavioral/emotional control, anxiety, positive affect, and an index of overall psychological distress/well-being) [34-43]. These measures were developed to assess two underlying health constructs, physical and mental health [44,45]. Most of the measures are from the MOS Functioning and Well-being Profile Core Subset (MOSFWBP-C) [46]; six of the measures are also part of the SF-36 which was derived from the MOSFWBP-C [41], and two are from the SF-20 [42]. The measures are summarized in the Appendix. Internal-consistency reliability of the baseline measures was acceptable for group comparisons (i.e. 0.70 or higher). Correlations between baseline and year two for each measure tended to be moderate (0.40–0.73, median 0.65), indicating considerable stability over time.

**Measures of disease and comorbidity**

Dichotomous indicators of whether the patient had hypertension, diabetes (Type I or II), myocardial infarction, congestive heart failure, and/or depressive symptoms or depressive disorder were constructed based on patient and doctor reports (see Sample and study design section).

Medical comorbidity was measured using an unweighted count of 15 chronic and acute conditions, defined from the health examination medical history interview. These conditions included chronic obstructive pulmonary disease, arthritis, anemia, asthma attacks in the past 6 months, irritable bowel, diverticulitis, hip im-
pairments, cancer, difficulty hearing, frequent sinus congestion, frequent cough, repeated diarrhea or constipation, varicose veins, and leg cramps. No attempt was made to weight the conditions, because unit weights are robust and tend to cross-validate about as well or better than non-unitary weights [47-49]. These 15 conditions were included on the basis of their prevalence and distinction from the four main chronic medical conditions or their complications. The comorbidity measure was available only on those who completed a medical history interview; those who did not were assigned the sample mean on this measure to maintain an adequate sample size. A dummy variable indicating whether or not the patient had these data available was included in the analysis to determine if those who received the medical history differed on the dependent measures from those who did not [50].

Analysis plan

Because we were concerned about selection effects (i.e. only individuals with mild disease exercise), we first examined the association between physical activity and the presence and severity of chronic illness [29]. We concluded that the variation in physical activity level was roughly similar across all chronic condition and severity groups, including depression, therefore enabling us to statistically control for the effects of disease and disease severity.

Regression methods were used to estimate the association of baseline physical activity with functioning and well-being (health) measures at baseline and 2 years after baseline. For each combination of physical activity and outcome measures, three types of regression analyses were performed:

1. **Cross-sectional**—health status at baseline was regressed on physical activity at baseline, controlling for all covariates (to be described below);
2. **Endpoint**—health status two years after baseline was regressed on physical activity at baseline, controlling for covariates;
3. **Change score**—change in health between baseline and two years was regressed on baseline levels of physical activity, controlling for all covariates.

For each of the three basic types of regression analyses, covariates included sociodemographic variables (age, education, gender, ethnicity (nonwhite)), an indicator of poverty level income (income below 200% of poverty level), study site, method of payment (prepaid vs fee-for-service), disease condition (hypertension, insulin-dependent diabetes, non-insulin dependent diabetes, congestive heart failure, myocardial infarction, subthreshold depression, current depressive disorder), comorbidity, seasonality, provider type (mental health vs medical provider), and other health behaviors (dummy variables for current smoker, past smoker, problem drinker, overweight) as well as general adherence to medical recommendations.

To allow for possible nonlinear effects of age [51] we included three dummy variables: 17-44, 45-64, and 65-74 (75+ age group was the holdout). Because the depressive symptoms indicator was included in all models, estimates of the relationship of physical activity and psychological distress/well-being are conservative and may be understated. That is, variation in depression is largely absorbed by the depressive symptoms covariate. In addition to the above, a series of 2-way interaction terms were included for each condition (i.e. condition by physical activity) to determine whether the effects of physical activity were different for those with vs without various conditions (e.g. depressive symptoms, congestive heart failure). Finally, a series of 3-way interaction terms was included, adding specialty type to each 2-way interaction, to determine if any differences in the relationship between activity and condition varied by specialty type (mental health vs medical provider). The statistical significance of the overall set of interaction terms was evaluated for each model.

Level of functioning and well-being is reported for each domain at baseline of the MOS study and 2 years later. The baseline cross-sectional data provide a point-in-time picture at the beginning of the study, and the endpoint data provide a similar snapshot 2 years later. We did not control for baseline status in the endpoint model (i.e. ANCOVA adjustment), because this variable may be correlated with the error term of the endpoint, a violation of statistical assumptions. In the presence of initial differences among compared groups, ANCOVA adjustment for baseline level of the outcome measure often leads to biased estimates of regression parameters [52-54]. We also conducted analyses using the difference between baseline and endpoint status as the dependent variable. Difference scores in functioning and well-being outcomes provide appropriately conservative
tests of null hypotheses and allow each individual to serve as his or her own control [55, 56]. We were unable to evaluate the association between changes in physical activity and changes in health outcomes because physical activity was measured only at baseline.

To evaluate the magnitude of any observed associations between physical activity and functioning and well-being, we obtained predicted health scores (least squares means) for those at a low level of physical activity (one standard deviation below the mean) and those at a high level of physical activity (one standard deviation above the mean) for each measure of physical activity. These were calculated separately for each chronic condition group in the medical specialty sector; separate estimates were obtained for patients of mental health providers for subthreshold depression and current depression (disorder).

To the extent that functioning and well-being outcomes for patients of a given provider are correlated, the effective sample size is reduced, requiring adjustment of standard errors and inference statistics. Therefore, we adjusted our models for these cluster effects using the Huber method [57].

Three weights were applied to the results which collectively adjust the analytic sample to better represent the original population of patients receiving care in each practice. The first weight compensates for the greater probability of selecting patients of clinicians with longer screening periods. The second weight compensates for the greater probability of selecting patients who visit their doctor frequently. The final weight compensates for the problem of attrition of the sample over time. In addition, we weighted the data for the probability of selection of the MOS clinicians, to make results more generalizable to the practices studied [26].

Regression coefficients were considered significant if the probability was less than or equal to 0.05. A more conservative Bonferroni-type correction was not considered appropriate because this would tend to overcorrect individual comparisons (Type II error), leading to insensitivity to true effects [58, 59]. Because the number of significant effects overall is inflated by chance, we interpret our results with this in mind.

RESULTS

As shown in Table 1, patients remaining in the analytic sample at 2 years differed significantly from those who were excluded. Reasons for being omitted from the analysis were dropout from the study (53%), missing data (29%), relocation (10%), and death (8%). Those who remained in the analytic sample were more likely to be white, have higher income, and...
Table 2. Summary of significant physical activity coefficients in multiple regressions with functioning and well-being dependent variables

<table>
<thead>
<tr>
<th>Functioning and well-being Measure</th>
<th>Perceived activity level</th>
<th>Total time spent exercising</th>
<th>Total time spent walking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline 2-year Change</td>
<td>Baseline 2-year Change</td>
<td>Baseline 2-year Change</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Role limitations—physical</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Role limitations—emotional</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Social activity limitations</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Energy/fatigue</td>
<td>-</td>
<td>p &lt; 0.05*</td>
<td>-</td>
</tr>
<tr>
<td>Pain severity</td>
<td>-</td>
<td>p = 0.05*</td>
<td>-</td>
</tr>
<tr>
<td>Sleep problems</td>
<td>-</td>
<td>p &lt; 0.05*</td>
<td>-</td>
</tr>
<tr>
<td>Psychological distress/well-being</td>
<td>-</td>
<td>p &lt; 0.05*</td>
<td>-</td>
</tr>
<tr>
<td>Depression</td>
<td>-</td>
<td>p &lt; 0.05*</td>
<td>-</td>
</tr>
<tr>
<td>Anxiety</td>
<td>-</td>
<td>p &lt; 0.001* p &lt; 0.05*</td>
<td>-</td>
</tr>
<tr>
<td>Positive affect</td>
<td>-</td>
<td>p &lt; 0.05*</td>
<td>-</td>
</tr>
<tr>
<td>Health distress</td>
<td>-</td>
<td>p &lt; 0.05* p &lt; 0.01*</td>
<td>-</td>
</tr>
<tr>
<td>Current health</td>
<td>-</td>
<td>p &lt; 0.01* p &lt; 0.05*</td>
<td>-</td>
</tr>
</tbody>
</table>

*p Overall test of all 2-way and 3-way interaction terms between chronic condition, physical activity, and provider specialty significant (p < 0.05).

more education, and to be older than the complete baseline sample. With respect to lifestyle characteristics, those who remained in the analytic sample were more physically active, less likely to be current smokers, more likely to have never smoked, and less likely to be problem drinkers. These differences, while statistically significant, are small in magnitude. For example, there are about 7% fewer nonwhites, 7% fewer current smokers, and 3% fewer problem drinkers at the 2-year follow-up than at baseline; those remaining at 2 years tended to be about 2 years older. Because of this and because of the weight included in the analyses for form completion as noted above, bias due to attrition is unlikely to unduly influence the conclusions.

Regression analyses examining the association between the three physical activity measures and the various functioning and well-being measures are summarized in Table 2. Each row represents one health measure and the three main columns represent the three selected physical activity measures. For each activity measure, three models are specified: baseline, 2-year endpoint, and change scores. Entries in the table are the significance of the coefficient for the physical activity measure in predicting the health outcome, controlling for all other variables in the model. An asterisk indicates those models in which the 2-way and 3-way interactions between the physical activity measure and all chronic conditions were significant (overall interaction test).

All but one of the significant coefficients were in the direction indicating that more physical activity was associated with better health; the exception was the change in physical functioning associated with baseline total time spent exercising, which was in the opposite direction. The majority of significant associations were observed for the total time spent exercising measure. At baseline, this measure was most often associated with measures that tend to be more indicative of physical health (e.g., physical, role, and social functioning, current health perceptions, pain); at 2 years, associations with mental health indicators were more common (e.g., sleep problems, psychological distress/well-being, health distress). Change scores were significant in only a few instances, as might be expected given the conservative nature of this analysis. The percent of variance explained ($R^2$) of the baseline models for the amount of time exercising measure ranged from 0.17 to 0.45 (median 0.29); for the endpoint models, the percent of variance explained ranged from 0.19 to 0.40 with a median of 0.26. The health behavior variables (exercise, smoking, overweight) added about 2% of additional variance explained beyond that accounted for by the other independent variables.

Table 3 presents predicted functioning and well-being scores (least squares means) for those with low and high physical activity levels at baseline for the total time spent exercising measure at the 2 year point. In these tests, differences between low and high activity levels are tested within each condition, and for the two depression conditions, within specialty type. We present this particular set of scores as illustrative of the other models. Interaction terms between particular conditions and physical activity were
<table>
<thead>
<tr>
<th>Measure</th>
<th>Baseline activity level</th>
<th>Hypertension</th>
<th>CHF</th>
<th>MI</th>
<th>Type I diabetes</th>
<th>Type II diabetes</th>
<th>Subthreshold depression</th>
<th>Current depression</th>
<th>Subthreshold depression</th>
<th>Current depression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>High</td>
<td>84.3</td>
<td>72.2</td>
<td>84.5</td>
<td>83.3</td>
<td>72.2</td>
<td>72.3</td>
<td>73.2</td>
<td>78.6</td>
<td>80.8</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>79.5</td>
<td>63.8</td>
<td>79.8</td>
<td>78.6</td>
<td>74.7</td>
<td>69.2</td>
<td>64.4</td>
<td>80.2</td>
<td>72.7</td>
</tr>
<tr>
<td>Role limitations due to physical health</td>
<td>High</td>
<td>75.6</td>
<td>64.5</td>
<td>77.1</td>
<td>62.2</td>
<td>70.8</td>
<td>71.7</td>
<td>41.6</td>
<td>51.8</td>
<td>66.9</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>71.6</td>
<td>50.6</td>
<td>50.5</td>
<td>63.3</td>
<td>68.0</td>
<td>63.5</td>
<td>35.9</td>
<td>55.1</td>
<td>53.1</td>
</tr>
<tr>
<td>Role limitations due to emotional problems</td>
<td>High</td>
<td>87.1</td>
<td>88.0</td>
<td>86.1</td>
<td>96.8</td>
<td>90.6</td>
<td>67.4</td>
<td>40.8</td>
<td>49.3</td>
<td>58.1</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>83.8</td>
<td>76.2</td>
<td>83.5</td>
<td>88.5</td>
<td>79.4</td>
<td>62.4</td>
<td>44.2</td>
<td>53.6</td>
<td>50.3</td>
</tr>
<tr>
<td>Social activity</td>
<td>High</td>
<td>91.6</td>
<td>91.3</td>
<td>98.8</td>
<td>88.4</td>
<td>89.1</td>
<td>84.4</td>
<td>74.9</td>
<td>73.2</td>
<td>68.4</td>
</tr>
<tr>
<td>limitations due to health</td>
<td>Low</td>
<td>90.5</td>
<td>86.2</td>
<td>88.7</td>
<td>87.1</td>
<td>86.7</td>
<td>80.3</td>
<td>69.6</td>
<td>74.7</td>
<td>67.3</td>
</tr>
<tr>
<td>Energy/fatigue</td>
<td>High</td>
<td>68.8</td>
<td>61.1</td>
<td>63.9</td>
<td>80.5</td>
<td>60.6</td>
<td>58.9</td>
<td>32.2</td>
<td>54.2</td>
<td>39.1</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>64.5</td>
<td>52.4</td>
<td>60.5</td>
<td>69.0</td>
<td>57.9</td>
<td>54.0</td>
<td>48.5</td>
<td>58.4</td>
<td>54.5</td>
</tr>
<tr>
<td>Pain severity</td>
<td>High</td>
<td>71.7</td>
<td>71.0</td>
<td>77.6</td>
<td>53.0</td>
<td>72.6</td>
<td>68.1</td>
<td>62.6</td>
<td>61.2</td>
<td>66.6</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>70.7</td>
<td>68.7</td>
<td>78.6</td>
<td>63.8</td>
<td>67.0</td>
<td>65.9</td>
<td>58.6</td>
<td>63.5</td>
<td>62.3</td>
</tr>
<tr>
<td>Sleep anxiety</td>
<td>High</td>
<td>80.8</td>
<td>73.4</td>
<td>82.4</td>
<td>82.9</td>
<td>80.6</td>
<td>68.4</td>
<td>65.2</td>
<td>70.2</td>
<td>65.4</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>78.7</td>
<td>71.1</td>
<td>75.1</td>
<td>78.0</td>
<td>78.6</td>
<td>67.1</td>
<td>59.4</td>
<td>70.6</td>
<td>64.6</td>
</tr>
<tr>
<td>Psychological distress/well being</td>
<td>High</td>
<td>82.6</td>
<td>84.9</td>
<td>84.8</td>
<td>90.5</td>
<td>81.1</td>
<td>68.9</td>
<td>57.9</td>
<td>65.5</td>
<td>57.4</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>80.7</td>
<td>82.0</td>
<td>79.6</td>
<td>80.2</td>
<td>80.0</td>
<td>67.6</td>
<td>54.4</td>
<td>62.7</td>
<td>55.3</td>
</tr>
<tr>
<td>Anxity</td>
<td>High</td>
<td>86.0</td>
<td>89.7</td>
<td>88.4</td>
<td>98.7</td>
<td>86.7</td>
<td>74.6</td>
<td>62.2</td>
<td>65.8</td>
<td>68.4</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>83.1</td>
<td>84.6</td>
<td>82.7</td>
<td>85.0</td>
<td>82.9</td>
<td>71.3</td>
<td>57.8</td>
<td>67.5</td>
<td>63.5</td>
</tr>
<tr>
<td>Positive affect</td>
<td>High</td>
<td>70.0</td>
<td>69.6</td>
<td>72.2</td>
<td>68.4</td>
<td>66.2</td>
<td>53.8</td>
<td>43.3</td>
<td>51.0</td>
<td>41.0</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>67.0</td>
<td>67.2</td>
<td>66.7</td>
<td>62.9</td>
<td>66.6</td>
<td>52.9</td>
<td>41.2</td>
<td>49.2</td>
<td>40.5</td>
</tr>
<tr>
<td>Depression</td>
<td>High</td>
<td>87.5</td>
<td>90.3</td>
<td>89.9</td>
<td>90.5</td>
<td>85.7</td>
<td>76.8</td>
<td>62.5</td>
<td>71.8</td>
<td>64.1</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>86.7</td>
<td>87.3</td>
<td>84.8</td>
<td>83.4</td>
<td>84.4</td>
<td>74.2</td>
<td>60.2</td>
<td>69.1</td>
<td>60.9</td>
</tr>
<tr>
<td>Health distress</td>
<td>High</td>
<td>91.3</td>
<td>90.9</td>
<td>94.6</td>
<td>87.2</td>
<td>88.9</td>
<td>84.8</td>
<td>76.6</td>
<td>74.1</td>
<td>83.4</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>90.1</td>
<td>85.6</td>
<td>87.3</td>
<td>87.7</td>
<td>84.5</td>
<td>79.8</td>
<td>66.4</td>
<td>79.2</td>
<td>75.5</td>
</tr>
<tr>
<td>Current health perceptions</td>
<td>High</td>
<td>72.0</td>
<td>57.7</td>
<td>72.8</td>
<td>57.2</td>
<td>65.9</td>
<td>64.5</td>
<td>53.5</td>
<td>67.8</td>
<td>63.1</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>67.7</td>
<td>52.7</td>
<td>67.0</td>
<td>55.4</td>
<td>59.6</td>
<td>57.5</td>
<td>47.4</td>
<td>65.8</td>
<td>56.3</td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01.

N ranges from 1568 to 1750.
All scores range from 0-100 with high score indicating better health.
often statistically significant, suggesting that the effect of the physical activity was different for that condition. Tests conducted of the difference in health between low and high physical activity levels within each condition were often significant, even for health measures in which the effect of physical activity for the overall sample was not significant. For example, differences in physical functioning between the low and high activity levels were significant for six of the nine chronic condition (and provider) groups (Table 3), even though the main effect of total time spent exercising was not significant (Table 2). Thus, the association of physical activity with health is highly dependent on the condition the patient has. The magnitude of the difference in scores for those groups and measures in which the difference was significant ranged between 0.04 and 0.62 of the baseline standard deviation; the majority were between 0.17 and 0.39, corresponding to the conventional small to medium effect size [60].

**COMMENTS**

This is the first time that naturally occurring levels of physical activity have been evaluated in relation to such a broad array of functioning and well-being outcomes over time in a large sample of patients with different chronic conditions. The number of 2-year functioning and well-being outcomes that were associated with baseline physical activity levels is notable, given the length of follow-up and the fact that chronic conditions, comorbidity, and other health behaviors were controlled for, all of which are associated with these outcomes. At baseline, the outcome measures that were more likely to be related to physical activity tended to be those that are more indicative of physical health (e.g. physical functioning, role limitations due to physical health, energy/fatigue, pain, current health perceptions) than of mental health [45]. However, many of the psychological distress/well-being measures were significant at 2 years even when they were not significant at baseline. This was in part due to the fact that the baseline indicator of the presence of depressive symptoms accounted for a considerable amount of cross-sectional variation in psychological distress/well-being and by 2 years, the association was less strong. These results are in contrast to a study of healthy adults aged 50–65 years, using similar MOS measures in which individuals who participated to a greater extent in an exercise program over a 1 year period had better scores on measures of physical health at the end of the year but no association with measures of mental health [61]. It is possible that the relationship of exercise to physical and mental health varies depending upon the extent of chronic conditions present.

Because the MOS is an observational study of naturally occurring levels of physical activity, causal inferences cannot be made definitely. One of the most likely alternative explanations for the observed associations is that people who exercise or are physically active are healthier to begin with and our control measures of disease and comorbidity do not fully capture these differences. A causal interpretation is strengthened, however, by our use of measures of disease and comorbidity verified by a health examination. Nevertheless, the possibility that other unmeasured patient characteristics were not controlled for tempers the strength of the inferences we can draw from these data [52].

One interpretation of the general lack of significant changes in functioning and well-being associated with baseline levels of physical activity is that if physical activity is causally associated with health benefits, the benefits had already occurred by baseline of the study and persisted over 2 years. Thus, few additional changes were likely to be observed.

Although our results suggest that physical activity may be associated with better health outcomes over time, what remains to be addressed is whether changes in physical activity are associated with changes in health outcomes. Paffenbarger and colleagues [5] present evidence that changes in physical activity are associated with changes in the probability of premature mortality, based on longitudinal data from an observational study of college alumni. Similarly, data from the Alameda County Study suggest that changes in physical activity over a 9-year period are associated with changes in risk of death [62]. It would be worthwhile to determine if an intervention program to increase levels of physical activity would lead to changes in these functioning and well-being outcomes.

The fact that few significant results were observed for the total time spent walking measure may indicate that milder forms of physical activity such as walking may be less useful in terms of functioning and well-being outcomes. The MOS did not measure the distinction between walking for exercise and walking in general (e.g. the MOS measure of walking
asked about walking more than two blocks either for exercise or to get somewhere). Thus, we are unable to directly address this question. More research is needed regarding how different types and intensities of exercise are associated with functioning and well-being outcomes.

We have tried to isolate the association between physical activity and functioning and well-being outcomes. However, health behaviors often cluster so that people who exercise tend to weigh less, eat better, and are less likely to smoke and drink alcohol [63]. Although these were controlled for in our analyses, an important future step will be to focus on the unique and combined effects of these health behaviors in explaining these outcomes.

The analyses are all based on self-reported measures of physical activity levels, and no physical fitness measures are available which could be used to validate them. However, other studies have shown that self-reports of physical activity levels are fairly accurate [64, 65]. Further, the MOS assessed physical activity levels only at baseline, thus we cannot know whether these levels of activity were maintained over time.

Given the difficulty of treating patients with chronic disease, and the focus of treatment on improving functioning and well-being, our research suggests that exercise may improve health for the chronically ill. Future studies can examine whether exercise interventions for chronically ill patients improve these types of outcomes.

Acknowledgements—This work was directly supported by a grant from the New England Medical Center, Inc. through the auspices of the Henry J. Kaiser Family Foundation. The Medical Outcomes Study (MOS) was supported by grants from the Robert Wood Johnson Foundation (RWJ), the Henry J. Kaiser Family Foundation (HKF), the Pew Charitable Trusts (PCT), the Agency for Health Care Policy and Research, and RAND. Individuals at these organizations who provided invaluable assistance include: Linda Aiken and Alan Cohen (RWJ), Barbara Kehrer and Alvin Tarlov (HKF), and Rebecca Rimal and Roseanne Siegel (PCT). Cooperation and assistance in recruiting physicians and patients were provided by the American College of Physicians, the American Academy of Family Physicians, the American Academy of Pediatrics, the American Psychiatric Association, and the American Psychological Association and their local chapters in the MOS sites.

We gratefully acknowledge the assistance of Barbara Genovese with data tabulation. We appreciate the contribution of MOS staff: Sandra Berry, Maureen Carney, Sherrie Kaplan, Elizabeth McGlynn, Lisa Meredith, Judith Pertman, Cathy Sherbourne, and John E. Ware, Jr. We also acknowledge the cooperation of the participating health maintenance organizations: the ANCHOR Organization of Health Maintenance in Chicago (Charles Hertz and William Rollow), CIGNA Healthplans of Southern California (B. Ross Landess), and the Harvard Community Health Plan in Boston (Stephen Schoenbaum). The conclusions are those of the authors and do not necessarily reflect the opinion of the sponsors or our institutions.

REFERENCES


57. Huber PJ. The Behavior of Maximum Likelihood Estimates under Various Conditions. *Fifth Berkeley


Appendix overleaf
APPENDIX

Summary of MOSFWBP-C measures of functioning and well-being administered at baseline and 2-years, mean scores at baseline and 2-years, and correlations between baseline and 2-years (N = 1568-2157)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Definition</th>
<th>No. of items</th>
<th>Reliability</th>
<th>Baseline mean, SD</th>
<th>2-Year mean, SD</th>
<th>2-Year stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>Extent to which health limits physical activities such as self-care, walking, climbing hills and stairs, bending, lifting, and moderate and vigorous activities (SF36) [34, 41]</td>
<td>10</td>
<td>0.93</td>
<td>71.13, 27.10</td>
<td>70.95, 27.00</td>
<td>0.73</td>
</tr>
<tr>
<td>Role limitations due to physical health</td>
<td>Limitations in work or other daily activities due to physical health during past 4 weeks including cut down amount of time, accomplished less than wanted, limited in kind of work, had difficulty performing activities (SF36) [35, 41]</td>
<td>4</td>
<td>0.87</td>
<td>53.74, 40.52</td>
<td>58.62, 42.42</td>
<td>0.51</td>
</tr>
<tr>
<td>Role limitations due to emotional problems</td>
<td>Limitations in work or other daily activities due to emotional problems such as feeling depressed during past 4 weeks including cut down amount of time, accomplished less than wanted, not doing as carefully as usual (SF36) [35, 41]</td>
<td>3</td>
<td>0.84</td>
<td>66.41, 40.36</td>
<td>72.45, 38.36</td>
<td>0.40</td>
</tr>
<tr>
<td>Social activity limitations due to health</td>
<td>Limitations in normal social activities during past 4 weeks due to physical health or emotional problems (SF36) [37, 41]</td>
<td>2</td>
<td>0.83</td>
<td>78.97, 25.30</td>
<td>80.92, 25.01</td>
<td>0.51</td>
</tr>
<tr>
<td>Energy/fatigue</td>
<td>Amount of time in past month felt energetic, full of pep, worn out, tired (SF36*) [40, 41]</td>
<td>4</td>
<td>0.85</td>
<td>55.54, 22.18</td>
<td>57.33, 21.70</td>
<td>0.62</td>
</tr>
<tr>
<td>Pain severity</td>
<td>Intensity of pain during past 4 weeks (SF20) [36, 42]</td>
<td>1</td>
<td>na</td>
<td>65.73</td>
<td>66.06</td>
<td>0.52</td>
</tr>
<tr>
<td>Sleep problems index</td>
<td>Sleep disturbance, adequacy, somnolence, and awaken short of breath during past 4 weeks [39]</td>
<td>6</td>
<td>0.79</td>
<td>71.12</td>
<td>73.60</td>
<td>0.65</td>
</tr>
<tr>
<td>Mental health index III</td>
<td>Summary index of depression, anxiety, positive affect during past month (SF36, SF20) [38, 41, 42]</td>
<td>5</td>
<td>0.95</td>
<td>70.49</td>
<td>73.19</td>
<td>0.69</td>
</tr>
<tr>
<td>Depression/behavioral-emotional control II</td>
<td>Amount of time in past month felt in low spirits, downhearted, depressed, moody, down in dumps, nothing to look forward to, not in firm control of behavior, not emotionally stable [38]</td>
<td>8</td>
<td>0.92</td>
<td>77.20</td>
<td>79.37</td>
<td>0.66</td>
</tr>
<tr>
<td>Anxiety II</td>
<td>Amount of time in past month very nervous person, tense, high strung, restless, or fidgety [38]</td>
<td>3</td>
<td>0.84</td>
<td>73.26</td>
<td>75.86</td>
<td>0.65</td>
</tr>
<tr>
<td>Positive affect II</td>
<td>Amount of time in past month been a happy person, felt calm and peaceful, felt cheerful and lighthearted, daily life interesting [38]</td>
<td>4</td>
<td>0.88</td>
<td>59.25</td>
<td>59.75</td>
<td>0.68</td>
</tr>
<tr>
<td>Health distress</td>
<td>Amount of time in past month felt distressed about health (e.g. discouraged by health, worried about health, afraid about health) [40]</td>
<td>6</td>
<td>0.94</td>
<td>78.54</td>
<td>81.99</td>
<td>0.59</td>
</tr>
<tr>
<td>Current health perceptions</td>
<td>Perceived ratings of overall current health (e.g. I have been feeling bad lately, my health is excellent) (SF20) [40, 42]</td>
<td>5</td>
<td>0.88</td>
<td>54.93</td>
<td>57.95</td>
<td>0.69</td>
</tr>
</tbody>
</table>

*Only 3 of 4 items were administered at 2 years.
na = not only available for single-item measures.
All measures scored on 0-100 scales; high score indicates better health.