Background: Previous observational and interventional studies have suggested that regular physical exercise may be associated with reduced symptoms of depression. However, the extent to which exercise training may reduce depressive symptoms in older patients with major depressive disorder (MDD) has not been systematically evaluated.

Objective: To assess the effectiveness of an aerobic exercise program compared with standard medication (ie, antidepressants) for treatment of MDD in older patients, we conducted a 16-week randomized controlled trial.

Methods: One hundred fifty-six men and women with MDD (age, ≥50 years) were assigned randomly to a program of aerobic exercise, antidepressants (sertraline hydrochloride), or combined exercise and medication. Subjects underwent comprehensive evaluations of depression, including the presence and severity of MDD using Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria and Hamilton Rating Scale for Depression (HAM-D) and Beck Depression Inventory (BDI) scores before and after treatment. Secondary outcome measures included aerobic capacity, life satisfaction, self-esteem, anxiety, and dysfunctional cognitions.

Results: After 16 weeks of treatment, the groups did not differ statistically on HAM-D or BDI scores (P = .67); adjustment for baseline levels of depression yielded an essentially identical result. Growth curve models revealed that all groups exhibited statistically and clinically significant reductions on HAM-D and BDI scores. However, patients receiving medication alone exhibited the fastest initial response; among patients receiving combination therapy, those with less severe depressive symptoms initially showed a more rapid response than those with initially more severe depressive symptoms.

Conclusions: An exercise training program may be considered an alternative to antidepressants for treatment of depression in older persons. Although antidepressants may facilitate a more rapid initial therapeutic response than exercise, after 16 weeks of treatment exercise was equally effective in reducing depression among patients with MDD.

Arch Intern Med. 1999;159:2349-2356
SUBJECTS AND METHODS

SUBJECTS

Subjects were recruited through flyers, media advertisements, and letters sent to local physicians and mental health facilities. Respondents who appeared to meet the study inclusion criteria attended an initial screening interview with a trained clinical psychologist (W.E.C., S.H.) who determined the presence and severity of MDD using relevant parts of the Diagnostic Interview Schedule26 and the 17-item Hamilton Rating Scale for Depression (HAM-D).27 Subjects were considered eligible for the study if they met criteria of the Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition (DSM-IV)28 for MDD (depressed mood or loss of interest or pleasure, and at least 4 of the following symptoms: sleep disturbance, weight loss or change in appetite, psychomotor retardation or agitation, feelings of worthlessness or excessive guilt, impaired cognition or concentration, and recurrent thoughts of death) and received a severity score of at least 13 on the HAM-D.

Exclusion criteria included current antidepressant use, use of other medications that would preclude random assignment to drug or exercise treatment (eg, quinidine, metoprolol), current alcohol or substance abuse, medical contraindications to exercise (eg, significant orthopedic problems or cardiopulmonary disease that would prevent regular aerobic exercise), primary psychiatric diagnosis other than MDD (eg, bipolar disorder, psychosis), evidence of acute suicidal risk, psychotherapy initiated within the past year, and ongoing participation in regular aerobic exercise.

DEPRESSION MEASURES

The HAM-D29 is a 17-item clinical rating scale that was used to determine study eligibility and treatment outcome. Patients who received a HAM-D score within the range of 13 through 18 were considered to be mildly clinically depressed, whereas patients who obtained a HAM-D score of more than 18 were considered to be moderately to severely depressed. To evaluate interrater reliability, 10 randomly selected interviews were independently rated by 2 clinicians. The intraclass correlation for both raters was 0.96. Every effort was made to ensure that clinical raters were unaware of patients’ treatment group assignment after randomization. Patients were instructed not to reveal which treatment they received at the time of their posttreatment assessments. The clinical psychologists administering the Diagnostic Interview Schedule and HAM-D were not involved in the provision of any of the treatments or in the interim assessments.

The BDI30 is a 21-item self-report questionnaire consisting of symptoms and attitudes relating to depression, including items such as self-dislike, suicidal ideation, insomnia, and sadness. The items are summed with a range of 0 to 63; higher scores indicate greater depression. The BDI has been shown to be a valid and reliable measure of depression severity.31 A meta-analysis of the internal consistency of the BDI estimates yielded a mean coefficient α of .86 for psychiatric patients and a mean correlation of the BDI and the HAM-D of 0.73.32

ADDITIONAL PSYCHOMETRIC MEASURES

The following set of secondary self-report measures was administered before and at the end of the 16-week intervention: (1) State-Trait Anxiety Inventory,39 consisting of 20 items designed to assess acute anxiety (eg, “I feel anxious”). Agreement with each statement is rated on a 4-point scale, with 1 indicating not at all to 4, very much so. Higher scores indicate higher levels of anxiety. (2) Rosenberg Self-Esteem Scale,40 a 10-item scale that measures global self-esteem (eg, “I take a positive attitude toward myself”). Agreement with each statement was rated on a 5-point scale, with 1 indicating strongly agree to 5, strongly disagree. The items were summed such that higher numbers indicate higher self-esteem. (3) Life Satisfaction Index,41 a 20-item index that measures overall life satisfaction (eg, “As I grow older, things seem better than I thought they would be”). Agreement with each statement was rated on a 3-point scale, and the items were summed such that higher numbers indicate greater life satisfaction. (4) Dysfunctional Attitudes Scale,42 a 40-item self-report questionnaire designed to assess the irrational beliefs and faulty assumptions about reality hypothesized by Beck to determine emotional disturbance (eg, “If I fail at my work, then I am a failure as a person”). Agreement with each statement was rated on a 7-point scale, so that higher scores reflect less presence of dysfunctional beliefs.

AEROBIC CAPACITY

At baseline and at the conclusion of 4 months of treatment, participants underwent a symptom-limited graded exercise has examined the effects of exercise as a treatment for depression in a group of 30 older adults.21 Although patients reported a significant reduction in depressive symptoms measured by the Beck Depression Inventory (BDI) compared with wait-list controls (subjects on a waiting list for treatment who served as controls), the study had important limitations, including a small sample size, imprecise diagnosis of depression, brief (ie, 6-week) treatment period, lack of a standardized exercise regimen, and failure to document exercise training.

Other studies of the effects of exercise training on older, healthy, nondepressed adults suggest a number of potential psychological benefits, including improvements in cognitive function, mood, and sense of well-being.22-27 A previous study in our laboratory found lower levels of depressive symptoms among older men following 4 months of exercise training.28 Other studies of older patients with medical conditions also have shown improvements of psychosocial functioning as a consequence of exercise training.20-33 Although these studies are suggestive, to our knowledge the therapeutic effects of exercise on clinical depression have not been evaluated systematically.

Our study accomplishes this goal through a comparison of exercise treatment with the current standard intervention—antidepressant medication—in a group of older adults with MDD. We addressed the following specific questions: (1) How do the therapeutic effects of group exercise training compare with those of antidepressant (sertraline hydrochloride [Zoloft]) therapy? (2) Is there added benefit from combining both treatment modalities? (3) Do patients with different severity levels of depressive symptoms respond differentially to the respective treatment ap-
approaches? and (4) How do the treatments compare with respect to the rate of change of depressive symptoms?

**SAMPLE CHARACTERISTICS**

One hundred fifty-six patients (aged 50-77 years) met study criteria and were randomized into the study. The primary reasons for patient exclusion were a failure to meet the criteria for MDD, a HAM-D score of less than 13, or a history of bipolar disorder (Figure 1). Of the patients who met final entry criteria, 48 were randomly assigned to the medication condition, 53 to the exercise condition, and 55 to the combined condition. The treatment groups did not differ with respect to age, sex, level of education, marital status, income, ethnic composition, or history of recurrent depression (Table 1).

**ADHERENCE TO TREATMENT**

Thirty-two patients (20.5%) dropped out before completing the entire 16-week protocol. Dropout rates did not vary significantly across treatment conditions ($\chi^2 = 0.63; P = .73$); 7 patients (14.6%) in the medication condition, 14 (26.4%) in the exercise condition, and 11 (20.0%) in the combined condition did not complete the study. Six patients (2 from each group) dropped out of the study after randomization but before treatment was initiated owing to dissatisfaction with their group assignments, but were included in the data analysis following the intent-to-treat principle. Ten patients dropped out owing to medication adverse effects (eg, age-
Adherence to medication treatment was evaluated by pill count; patients appeared to take medications as prescribed (no patient deviated by >5% from the prescribed dosage). The median peak dose of sertraline hydrochloride was 100 mg for the medication and combined treatment groups. Attendance for the 2 exercise conditions was comparable, with patients in the exercise condition attending a median of 43 sessions (89.6% of scheduled sessions) and patients in the combined condition attending an average of 44 sessions (91.7% of the scheduled sessions). Similarly, the mean percentage of time in target heart rate training range was virtually identical for the exercise (82.4%) and combined (82.4%) groups.

CHANGES IN AEROBIC CAPACITY

The ANCOVA revealed that the treatment groups differed with respect to improved aerobic capacity (F2,153 = 5.62; P = .004) and treadmill test duration (F2,153 = 7.68; P < .001). Patients in the exercise and combination groups showed significant improvements in aerobic capacity (Figure 2), whereas patients in the medication group did not. Patients in the exercise condition achieved an 11% improvement in aerobic capacity of 2.7 ± 2.8 mL/kg per minute, and patients in the combined condition achieved a 9% improvement in aerobic capacity of 2.3 ± 2.2 mL/kg per minute; patients in the medication condition showed minimal (<3%) improvement. A similar pattern was observed with respect to exercise tolerance, with patients in the exercise group showing a 15% improvement in test duration, those in the combination group showing an 11% improvement, and those in the medication group showing a less than 3% improvement.

EFFECTS ON DEPRESSION

All 3 groups exhibited a significant decline in depressive symptoms. Figure 3 displays the mean HAM-D and BDI scores at study entry and at 16 weeks. The treatment groups did not differ significantly on baseline levels of depression as measured by the HAM-D (F2,153 = 0.96; P = .39) or the BDI (F2,153 = 0.90; P = .40). A 1-way multivariate analysis of variance revealed no statistically significant differences among the groups on the HAM-D or the BDI (Wilks’ λ1,44 = 0.98; P = .67). The ANCOVA models controlling for baseline levels of depression also did not yield any significant treatment effects with respect to the HAM-D (F2,152 = 0.61; P = .55) or BDI (F2,152 = 1.01; P = .37).

Finally, based on DSM-IV criteria for MDD36 and assuming that patients who dropped out of the study were still depressed, the percentage of patients who were no longer depressed at study entry did not differ significantly between the treatment groups. Approximately 82.4% of patients took medications (no patient deviated by >5% from the prescribed dosage). The median peak dose of sertraline hydrochloride was 100 mg for the medication and combined treatment groups. Attendance for the 2 exercise conditions was comparable, with patients in the exercise condition attending a median of 43 sessions (89.6% of scheduled sessions) and patients in the combined condition attending an average of 44 sessions (91.7% of the scheduled sessions). Similarly, the mean percentage of time in target heart rate training range was virtually identical for the exercise (82.4%) and combined (82.4%) groups.
classified as clinically depressed at the end of the 4-month treatment period did not differ across treatment groups ($\chi^2 = 0.79; P = .67$): 32 patients (60.4%) in the exercise group, 33 (68.8%) in the medication group, and 36 (65.5%) in the combination group no longer met DSM-IV criteria for MDD according to the clinician-rater who was unaware of patients’ group status. When the additional criteria of a HAM-D score greater than 6 was added to the DSM-IV classification, the groups again did not differ ($\chi^2 = 1.08; P = .58$), with 25 patients (47.2%) in the exercise group, 27 (56.2%) in the medication group, and 26 (47.3%) in the combination group classified as being no longer clinically depressed.

### RATE OF TREATMENT RESPONSE

The growth curve analysis for the HAM-D revealed that the rate of treatment response (depression score as a function of time in treatment) differed in a statistically significant manner ($P = .02$) across the treatment groups depending on the initial severity of depression. This relation can be seen best in the fitted growth curves presented in the left panel of Figure 4, which shows the shape of the change in depression scores at selected values of high and low initial depression. Depressed patients in the medication group exhibited a more rapid initial therapeutic response (within the first few weeks) compared with patients in the other groups. Also, mildly depressed patients appeared to respond more quickly to the combination of medication and exercise than did moderately to severely depressed patients. The trajectory of BDI scores (Figure 4, right) also varied as a function of treatment group and initial severity level of depression such that patients with mild depressive symptoms at baseline who received combination therapy exhibited a more rapid response compared with patients receiving only medication or only exercise. Summaries of the growth curve analyses for the HAM-D and BDI are given in Table 2 and Table 3, respectively.

### ADDITIONAL PSYCHOLOGICAL VARIABLES

The groups did not differ at baseline on self-reported levels of anxiety, self-esteem, life satisfaction, or dysfunctional attitudes (Table 4). At 16 weeks, all groups had improved on each of these measures, but these changes were not significantly different across groups.

### COMMENT

The results of this study provide empirical support for the notion that a group program of aerobic exercise is a feasible and effective treatment for depression in older...
Most patients were able to complete the exercise training protocol successfully. Dropout rates and adherence data compare favorably with those reported in other studies of exercise in older populations and suggest that the presence of clinical depression does not preclude participation in an exercise program. Moreover, exercising subjects achieved small but clinically and statistically significant improvements in aerobic capacity, which were comparable to changes observed in nonclinical populations of older adults.

Patients also appeared to achieve significant clinical improvement with exercise training. Among the 156 patients who entered the trial, 60.4% of patients in the exercise condition, 68.8% of patients in the medication condition, and 65.5% of patients in the combined condition no longer met DSM-IV criteria for MDD. All 3 groups also showed significant reductions on HAM-D and BDI scores, indicating that their clinical symptoms were significantly reduced. That exercise was equally effective as medication after 16 weeks of treatment is consistent with findings of other studies of exercise training in younger depressed adults. The magnitude of reductions in depression scores is also comparable to the levels achieved using sertraline in other clinical trials of depression. Moreover, the changes in depressive symptoms found for all treatments in our study are consistent with the extent of improvements reported in more than a dozen studies of psychosocial interventions for MDD. For example, in the National Institute of Mental Health Collaborative Depression Study, 36% of patients undergoing cognitive behavior therapy, 43% of patients undergoing interpersonal therapy, and 42% of patients receiving medication (imipramine hydrochloride) were considered “recovered,” compared with 47.2% undergoing exercise, 56.2% receiving medication (sertraline), and 47.3% receiving a combination of exercise and medication in our study.

Table 2. Solution for Growth Curve Model of HAM-D*

<table>
<thead>
<tr>
<th>Source</th>
<th>df†</th>
<th>F Statistic</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline HAM-D</td>
<td>1</td>
<td>18.56</td>
<td>.001</td>
</tr>
<tr>
<td>Early weeks</td>
<td>1</td>
<td>0.98</td>
<td>.32</td>
</tr>
<tr>
<td>Later weeks</td>
<td>1</td>
<td>1.18</td>
<td>.28</td>
</tr>
<tr>
<td>Treatment group</td>
<td>2</td>
<td>2.22</td>
<td>.11</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early weeks × group</td>
<td>2</td>
<td>4.69</td>
<td>.009</td>
</tr>
<tr>
<td>Later weeks × group</td>
<td>2</td>
<td>3.89</td>
<td>.02</td>
</tr>
<tr>
<td>Baseline HAM-D × early weeks</td>
<td>1</td>
<td>4.48</td>
<td>.04</td>
</tr>
<tr>
<td>Baseline HAM-D × later weeks</td>
<td>1</td>
<td>2.99</td>
<td>.08</td>
</tr>
<tr>
<td>Baseline HAM-D × group</td>
<td>2</td>
<td>2.73</td>
<td>.07</td>
</tr>
<tr>
<td>Baseline HAM-D × early weeks × group</td>
<td>2</td>
<td>4.63</td>
<td>.01</td>
</tr>
<tr>
<td>Baseline HAM-D × later weeks × group</td>
<td>2</td>
<td>3.98</td>
<td>.02</td>
</tr>
</tbody>
</table>

*HAM-D indicates Hamilton Rating Scale for Depression; early weeks, slope from week 1 to week 4; later weeks, slope from week 4 to week 16.†Denominator df = 984 for all tests.

Table 3. Solution for Growth Curve Model of BDI*

<table>
<thead>
<tr>
<th>Source</th>
<th>df†</th>
<th>F Statistic</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline BDI</td>
<td>1</td>
<td>29.72</td>
<td>.001</td>
</tr>
<tr>
<td>Early weeks</td>
<td>1</td>
<td>0.02</td>
<td>.89</td>
</tr>
<tr>
<td>Later weeks</td>
<td>1</td>
<td>0.00</td>
<td>.99</td>
</tr>
<tr>
<td>Treatment group</td>
<td>2</td>
<td>6.24</td>
<td>.002</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early weeks × group</td>
<td>2</td>
<td>4.65</td>
<td>.01</td>
</tr>
<tr>
<td>Later weeks × group</td>
<td>2</td>
<td>4.82</td>
<td>.008</td>
</tr>
<tr>
<td>Baseline BDI × early weeks</td>
<td>1</td>
<td>6.69</td>
<td>.01</td>
</tr>
<tr>
<td>Baseline BDI × later weeks</td>
<td>1</td>
<td>3.55</td>
<td>.06</td>
</tr>
<tr>
<td>Baseline BDI × group</td>
<td>2</td>
<td>8.98</td>
<td>.001</td>
</tr>
<tr>
<td>Baseline BDI × early weeks × group</td>
<td>2</td>
<td>5.79</td>
<td>.002</td>
</tr>
<tr>
<td>Baseline BDI × later weeks × group</td>
<td>2</td>
<td>6.10</td>
<td>.003</td>
</tr>
</tbody>
</table>

*BDI indicates Beck Depression Inventory; early weeks, slope from week 1 to week 4; later weeks, slope from week 4 to week 16.†Denominator df = 984 for all tests.
Although patients in all 3 groups achieved comparable improvement by the end of the 16-week treatment program, the rate of response appeared to vary as a function of depression severity. Those patients who received medication alone appeared to have the fastest response to treatment. Patients with less severe depression appeared to respond more quickly to the combination of exercise and medication than those more severely depressed counterparts.

The mechanisms responsible for the reduction in depressive symptoms are unknown. Several observational studies have shown an association between enhanced physical fitness and improved mental health. For example, in a secondary analysis of 4 surveys, positive affect was related directly to the amount of physical activity in household populations of the United States and Canada. Data from cross-sectional studies of active and sedentary individuals also suggest an association between physical activity and mood. A number of studies have found that exercise-trained patients obtain significantly lower depression scores on self-report measures than sedentary patients. Data from longitudinal studies, however, have provided mixed results, with no clear evidence that improved aerobic capacity is responsible for reduced depression. Data from our study revealed that, among patients who exercised, a small, statistically significant correlation \( r = -0.25; P = .02 \) was found between changes in aerobic capacity and changes in HAM-D scores, suggesting that improved aerobic capacity may have accounted for at least part of the reductions in depression. Other biological mechanisms, including alterations in central norepinephrine activity, reduced activity of the hypothalamopituitary-adrenocortical axis, and increased secretion of beta endorphins, have been suggested as mechanisms by which exercise improves mood, but no data were available to address these possibilities. Psychological mechanisms also may be responsible for the effects of exercise on mood, including increased feelings of self-efficacy, improved self-concept, and reduced dysfunctional or negative thought patterns. Our study offers no definitive conclusions about what psychological processes may mediate the improved mood associated with exercise training.

This was a short-term (16-week) intervention study. The positive changes that were observed among patients treated with exercise also were observed in those receiving medication only. It remains unknown whether differential outcomes might have been obtained with a longer duration of treatment. It also is not clear how resistant the treatment effects will be to relapse over time. Other studies have reported that relapse and long-term recurrence rates for MDD range from 50% to 80%. Although exercise and antidepressants may be comparable in their short-term effectiveness, there may be important differences in maintenance. Follow-up of these patients is ongoing, and the enduring effects of these interventions will be evaluated subsequently.

Patients also were sufficiently motivated to volunteer for a study of exercise training, tended to be highly educated, and were healthy enough to participate in an exercise program. The extent to which these findings may be generalizable to other older depressed individuals will need to be studied.

Another limitation of our study was the absence of a true no-treatment control group. This prevents us from specifying a “spontaneous recovery” rate against which our treatment results can be compared. After much consideration of this issue during the initial planning of the study, we decided for methodological and ethical reasons to assign patients only to the 3 active treatment groups. Our objective was to compare exercise treatment with an established treatment modality of known effectiveness. The additional value to be gained by including a no-treatment or placebo condition in our study was judged to be minimal compared with the potential negative clinical impact on those study participants assigned to an intervention known to be of lesser effectiveness. It has been estimated that approximately 30% of patients respond to placebo, for example, 21% of patients receiving placebo were considered recovered in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. Thus, patients undergoing exercise training in our study achieved approximately 2-fold greater rates of recovery from MDD compared with patients receiving placebo in previous clinical trials. It is therefore unlikely that the benefits of exercise can be attributed solely to placebo effects or spontaneous remission.

Also, because exercise was performed in a group setting, it is possible that the social interaction of study participants may have had a beneficial effect. Future research on exercise may need to control for the level of social involvement by examining the effects of the exercise setting (eg, home vs program-based) on response to treatment. Based on our results, it appears that exercise is associated with clinically significant improvements in depression that are comparable to those of antidepressants for the clinically depressed older adult when performed in a structured and supervised setting.

Accepted for publication March 16, 1999.

Supported by grants MH49679, HL43028, HL49572, and MO1-RR-30 from the National Institutes of Health.
REFERENCES


