

A Randomized Controlled Trial of High Versus Low Intensity Weight Training Versus General Practitioner Care for Clinical Depression in Older Adults

Nalin A. Singh,^{1,2} Theodora M. Stavrinou,² Yvonne Scarbek,³ Garry Galambos,³ Cas' Liber,³ and Maria A. Fiatarone Singh^{4,5}

¹Royal Prince Alfred Hospital, ²Balmain Hospital, and

³Department of Psychiatry, Rozelle Hospital, Central Sydney Health Service, New South Wales, Australia.

⁴Department of Exercise and Sports Science, University of Sydney, New South Wales, Australia.

⁵Senior Research Staff, Hebrew Rehabilitation Center for Aged, Boston, Massachusetts.

Background. Although exercise has been shown to relieve depression, little is known about its mechanism or dose-response characteristics. We hypothesized that high intensity progressive resistance training (PRT) would be more effective than either low intensity PRT or standard care by a general practitioner (GP) in depressed elderly persons, and that high intensity PRT would provide superior benefits in quality of life, sleep quality, and self-efficacy.

Methods. Sixty community-dwelling adults >60 years with major or minor depression were randomized to supervised high intensity PRT (80% maximum load) or low intensity PRT (20% maximum load) 3 days per week for 8 weeks, or GP care.

Results. A 50% reduction in the Hamilton Rating Scale of Depression score was achieved in 61% of the high intensity, 29% of the low intensity, and 21% of the GP care group ($p = .03$). Strength gain was directly associated with reduction in depressive symptoms ($r = 0.40$, $p = .004$), as was baseline social support network type ($F = 3.52$, $p = .015$), whereas personality type, self-efficacy, and locus of control were unrelated to the antidepressant effect. Vitality quality-of-life scale improved more in the high intensity group than in the others ($p = .04$). Sleep quality improved significantly in all participants ($p < .0001$), with the greatest relative change in high intensity PRT ($p = .05$).

Conclusions. High intensity PRT is more effective than is low intensity PRT or GP care for the treatment of older depressed patients.

DEPRESSION has been identified as one of the future key areas in the health of older adults (1), resulting in substantial personal and societal burden including role impairment, reduced quality of life, and excess health care expenditures. Clinically defined major depressive disorder is common, with a lifetime prevalence of 16% in the recent U.S. National Comorbidity Survey Replication sample (2). Among these, only 22% of cases reported in the previous year had received adequate treatment, and the vast majority of patients with chronic major depression are “misdiagnosed, receive inappropriate or inadequate treatment, or are given no treatment at all,” according to the National Depressive and Manic-Depressive Association Consensus Statement on the Undertreatment of Depression (3). Meta-analysis of trials of antidepressants in older adults indicates high dropout rates (approximately 25%) and rates of adverse events (approximately 60%) (4), during an average treatment duration of only 5 weeks. Therefore, the search for effective and less toxic treatments is a priority in research, particularly for older adults who may be less tolerant of antidepressant medications than are younger patients (5).

Exercise is one such potential alternative treatment, as both aerobic and resistance training have been shown to significantly improve depression in randomized controlled trials of clinically depressed older adults (6–8). In a previous randomized controlled trial we found progressive resistance

training (PRT) to have a clinically important antidepressant effect in a group of older patients with minor depression, dysthymia, and major depression (8), extending similar findings from younger depressed patients (9,10). Because our control group received an attention control activity, we could not compare the efficacy of exercise directly with that of standard care. Despite this limitation, we found two interesting results that are the basis for our further hypothesis testing in this study:

1. We found significant effects on major depression but not minor depression. This finding was contrary to a widely held belief that exercise is suitable for minor depression but not major depression.
2. Within the exercise group, the higher the relative intensity of the training, the greater the reduction in depression, suggesting that there may be a dose-response relationship in exercise treatment for depression.

Therefore, we designed the current study to test three hypotheses: 1) PRT exercise is an effective antidepressant in older adults with clinical depression; 2) high intensity PRT is superior to low intensity exercise; and 3) high intensity PRT is superior to standard general practitioner (GP) care of depression in older adults.

The secondary aims of the trial were to compare the effects of the three interventions on measures of quality of

life, sleep disturbance, and self-esteem, all of which have been shown to improve in other antidepressant trials (11,12). In addition, we sought to explore potential relationships between baseline participant characteristics such as outcome expectation, social network, self-esteem, and locus of control, and treatment response. These factors have been previously related to antidepressant medication efficacy, behavioral change, or exercise adaptations, and we postulated that they might in part explain the benefit of exercise as an antidepressant.

METHODS

Study Design

This was a randomized controlled 8-week study comparing the effects of assignment to one of three interventions for clinical depression: two exercise interventions (high intensity PRT [HIGH] and low intensity PRT [LOW]) and a usual care group getting standard GP care. Randomization followed baseline assessment and was done by a computer-generated random number permutation program in blocks of 15. Randomization lists were prepared by a statistician uninvolved in participant assessments or interventions, and were concealed in opaque sealed envelopes from investigators and participants until the time of allocation. The ethics committee of the Central Sydney Area Health Service approved the study, and all study participants provided written informed consent.

Study Population

Participants were recruited through 42 individual GPs in the Central Sydney area. Community-dwelling patients over 60 years of age received a letter and a Geriatric Depression Scale (GDS) (13) to complete if interested. Participants were blind to the investigators' hypotheses. Participants included in the study were aged >60 years; fulfilled DSM-IV (Diagnostic and Statistical Manual of Mental Disorders) (14) diagnostic criteria for major depression, minor depression, or dysthymia; and had a GDS score of ≥ 14 . Participants were excluded if demented clinically according to DSM-IV criteria or if their Folstein Mini-Mental State Examination (15) score was ≤ 23 , if they were suffering from unstable medical disease which would preclude resistance training, had bipolar disorder or active psychosis, or were determined by the study physician to be actively suicidal. They were also excluded if they were currently seeing a psychiatrist, prescribed antidepressant drugs within the last 3 months, or were currently participating in any exercise training more than twice a week.

Intervention

Participants assigned to HIGH underwent a regimen of supervised high intensity PRT of the large muscle groups, 3 days per week for 8 weeks. Exercise machines included chest press, upright row, shoulder press, leg press, knee extension, and knee flexion (Keiser Sports Health Equipment, Fresno, CA). For the HIGH group the resistance was set at 80% of the one repetition maximum (1RM; the maximal load that could be lifted fully one time only) (16) on each machine. To maintain the intensity of the stimulus, the load was increased

at each session as assessed by the Borg scale of perceived exertion (17), by keeping the rating between 15 and 18 on the 20-point scale. Strength testing was repeated every 4 weeks to establish a new baseline value. On each machine, participants performed 3 sets of 8 repetitions. The LOW group underwent low intensity resistance using the exact same regimen, except they were trained at 20% 1RM and not progressed, although perceived exertion was assessed at each session. Each session lasted approximately 60 minutes and was followed by 5 minutes of stretching. Discussion of depression was minimized. All sessions occurred at Balmain Hospital in an outpatient gymnasium setting in groups of 1-8 participants.

Standard Care

All participants not randomly assigned to resistance training received usual care from their GP. After being randomized to GP care, participants were given a letter to take to their GP in the next week. The letter (a) stated that patients were enrolled in the study and (b) noted their GDS score with a code outlining severity. GPs were informed that the management of the patients' depression was entirely unrestricted. GPs of participants in the exercise groups also received a letter notifying them of group assignment and requesting that no new treatment be commenced for depression in the next 8 weeks. All participants were asked not to commence any new exercise regimen, other than what had been prescribed within the study.

Outcome Measures

A blinded psychiatrist performed all outcome measures at baseline and 8 weeks. To assess the success of blinding at follow-up, the psychiatrist was asked to guess to which group the patient had been randomized. Strength testing and adverse events collection was carried out by investigators who were not blinded to group assignment.

Depression.—Our cohort and outcome measures were defined precisely in concordance with clinical definitions of disease severity and response to treatment as utilized in clinical trials of antidepressant medications (18). DSM-IV symptoms and psychiatric diagnoses were assessed by structured clinician interview, according to the DSM-IV manual (14). The Hamilton Rating Scale of Depression (HRSD) (17 item scale, 0-52) (19), was the therapist-rated measure of depression (5,8,20), and as recommended, a clinical response was defined as a 50% or greater improvement in HRSD score. The primary self-rated measure of depression was the GDS, a well-validated tool in elderly people (13).

Outcome expectation.—Each patient was asked at baseline (before randomization) to rate on a Likert scale from 0 to 10, where 10 was *strongly agree* and 0 was *strongly disagree*, the following two statements: (1) I am convinced that the exercise program will help me lift my depression and (2) I am convinced that my GP's treatment will help me lift my depression.

Personality.—The Eysenck Personality Questionnaire (EPQ) or EPQ-Revised short form was used (21).

Social network.—Social support network was assessed with the Wenger social support network instrument (22). This well validated typology categorizes older individuals as either local family-dependent, locally integrated, local self-contained, private restricted, or wider community-focused. Strong social support networks have been associated with better emotional health, morale, and quality of life, and with reduced use of health care services (23).

Self-efficacy.—A general well validated measure was used, the Self-Efficacy Scale of Sherer (24). The items focus on three areas: 1) willingness to initiate behavior, 2) willingness to expend effort in completing the behavior, and 3) persistence in the face of adversity.

Locus of control.—The validated Multidimensional Health Locus of Control (25) was used to assess this domain. This scale provides a measure of three dimensions of health locus of control: Internality, Powerful Others, and Chance.

Quality of life.—Health-related quality of life was measured by the Medical Outcomes Survey (Short Form 36), a widely used tool in older adults and intervention trials which is well validated and sensitive to change (26).

Sleep.—The Pittsburgh Sleep Quality Index is a well validated self-rated questionnaire assessing sleep quality over a 1-month time interval (27,28).

Physiological measures.—The IRM was used to determine muscle strength (16).

Monitoring

Participants were monitored each week by N.S. for adverse events or suicidality. The study protocol called for psychiatric referral if suicidal ideation occurred during the study.

Statistical Analysis

Sample size calculations were based on the ability to discern a treatment response of 60% in HIGH, and 30% in LOW and GP standard care groups, on the basis of our previous findings (8,12) and on average effect sizes (ES) reported in the literature for antidepressant medication or placebo (4,29). We estimated that we required 15 participants in each group at an alpha level of 0.05 and a beta of 0.10 (90% power) (30). Sample sizes were inflated by 25% to allow for possible attrition. All data were analyzed with Statview or SuperAnova statistical software (Abacus Concepts, Berkeley, CA). Continuous data are described as the mean (standard error) or median and range as appropriate for non-normally distributed data, except for quality-of-life measures, which are presented as mean (standard deviation) to allow comparison with reported normal values. Baseline differences in group characteristics were analyzed by analysis of variance (ANOVA) for continuous variables and chi-square or Fisher's exact test for categorical data. A repeated-measures ANOVA was used to analyze the effect of time, treatment, and time by treatment interactions for all continuous outcome variables

at baseline and at 8 weeks. Analysis of covariance models adjusting for baseline score were used to analyze relative changes over time in outcome measures between groups. Categorical outcomes were analyzed by logistic regression models with group assignment and appropriate baseline scores as independent predictors in the models. ES were calculated according to the method of Cohen for primary outcome variables as the mean change/standard deviation, and relative ES as the Experimental ES - Control (GP care) ES. Relationships between variables of interest were analyzed with regression models. A two-sided *p* value less than or equal to .05 was considered to indicate statistical significance.

RESULTS

Recruitment

Our recruitment flow chart is presented in Figure 1. The three major reasons for ineligibility after telephone screening were inability to make the time commitment, current antidepressant use, and nonclinical depression.

Participant Characteristics

Baseline participant characteristics are presented in Tables 1 and 2. There were no significant differences between groups in demographic variables, health status, psychological history, or any other measured characteristics other than marital status.

Dropouts

There were six dropouts (10%) from the study (two from HIGH, three from LOW, and one from GP care), none of whom agreed to return for final assessments. Reasons for dropout were failure to return after first session due to disinterest in exercise (two in HIGH, one in LOW), pain (two in LOW), and hospitalization (one control).

Compliance With Treatments

Compliance with exercise sessions (including dropouts) was 95%–100% in HIGH and 99%–100% in LOW (*p* = .90). Compliance in those who completed the study was 100% (range 12–100) in HIGH versus 100% (range 16–100) in LOW. Average intensity of training was $85 \pm 2.6\%$ in HIGH and $20 \pm 0.8\%$ in LOW, indicating successful maintenance of the intended dose of weight lifting exercise in the two experimental groups.

In the GP care group, all patients attended their GP's practice within 7 days of randomization. Treatment included antidepressant medication in 8 of 19 (42%), counseling alone in 2 of 19 (10%), and psychiatric referral in 1 of 19 (5%). Therefore, a total of 10 of 19 (52%) commenced some formal treatment for their depression. The average of five health care professional visits in the GP care group during the 8 weeks of the trial (see Table 3) were predominantly GP visits.

Adverse Events

Adverse events recorded weekly throughout the 8 weeks of the trial are shown in Table 3. There were no significant differences.

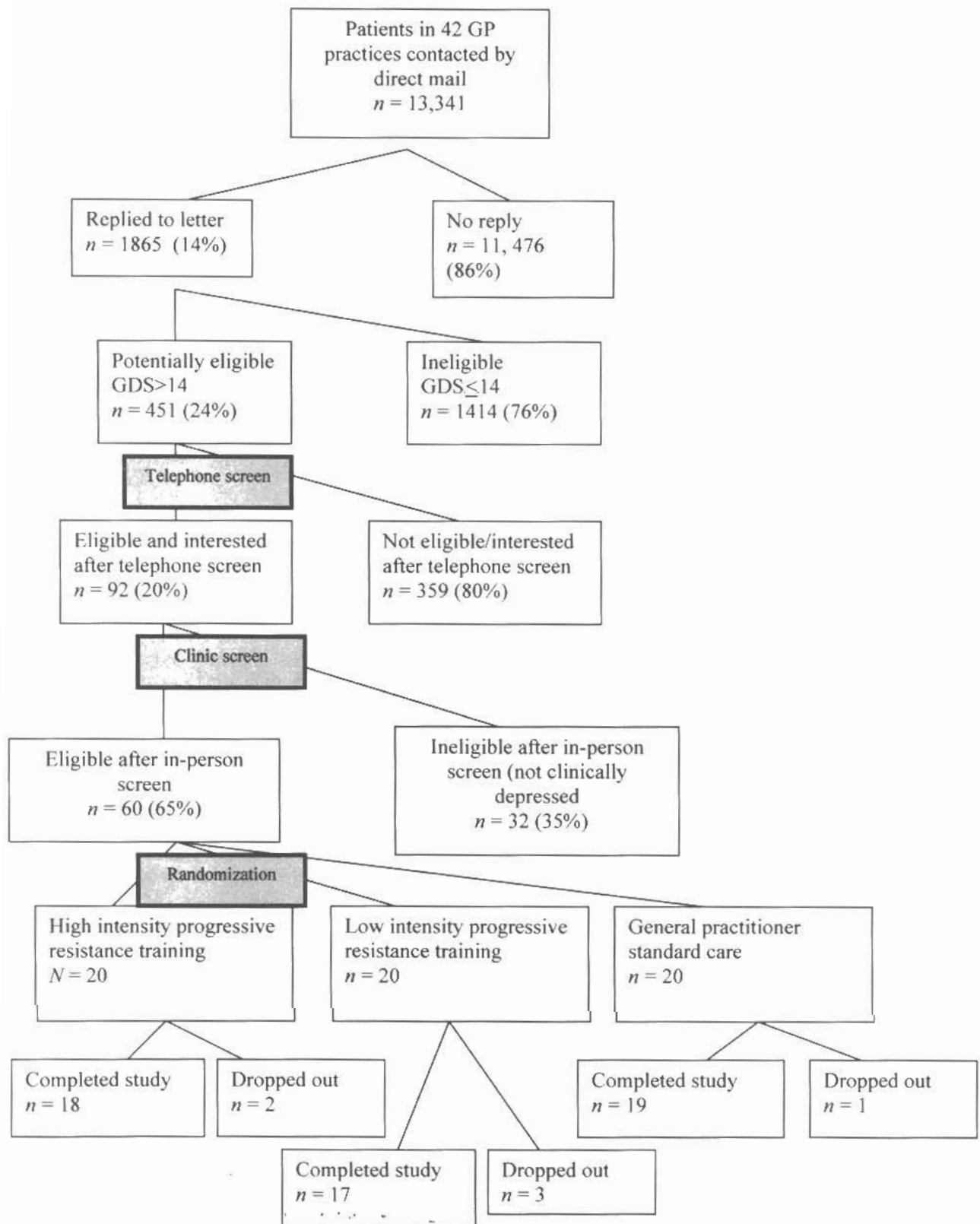


Figure 1. Flow of participants through the trial. GDS = Geriatric Depression Scale (0–30); normal score <9. See text for reasons for ineligibility and dropout. All dropouts were excluded from analysis of outcomes (refused to attend final assessment).

Table 1. Baseline Characteristics

Variable	HIGH (N = 20)	LOW (N = 20)	GP Care (N = 20)	p Value
Age y (range 60–85)	69 ± 5	70 ± 7	69 ± 7	.8
Marital status, n				.05
Married	9	2	13	
Widowed	4	7	2	
Divorced	6	8	4	
Separated	1	3	1	
Sex, n				.8
Men	9	8	10	
Women	11	12	10	
Residence, n				.9
House	14	13	13	
Apartment	6	7	7	
Annual income (AUD)				.5
<\$15,000	14	13	0	
\$15,000–\$30,000	5	7	8	
>\$30,000	1	0	12	
Smoking currently (n)	3	3	2	.8
Ethanol, g/wk	88 ± 125	72 ± 108	45 ± 58	.4
Education, y	11 ± 2	12 ± 3	11 ± 3	.7
Chronic diseases, n	3.7 ± 1.7	3.0 ± 1.0	3.0 ± 1.0	.1
Medications per day, n	4.6 ± 3.5	3.6 ± 3.3	4.2 ± 2.2	.6
Mini-Mental State Examination Score (0–30)	29 ± 1.3	29 ± 0.6	29 ± 1.2	.6

Note: HIGH = high intensity progressive resistance training group; LOW = low intensity progressive resistance training group; GP = general practitioner.

Adequacy of Blinding

Sixty-three percent of the time, the psychiatrist guessed group assignment incorrectly. When the guess was divided into two rather than three categories (exercise vs non-exercise), the blinded psychiatrist guessed incorrectly 50% of the time, suggesting successful maintenance of blinding.

Primary Outcomes

As shown in Table 4, there was a highly significant improvement over time in self-rated (GDS, $F = 92.20$, $p < .0001$) and therapist-rated (HRSD, $F = 74.87$, $p < .0001$) scales across all groups, but improvements were significantly larger in the HIGH group than in the other two groups for the GDS ($F = 5.58$, $p < .006$), with a similar direction of effect for the HRSD ($F = 1.98$, $p = .14$) in repeated measures ANOVA models of the absolute scores. HIGH patients had more than double the relative response (% change) of GP care participants in both GDS ($58 \pm 7\%$ vs $23 \pm 7\%$ reduction; $p = .001$) and HRSD scales ($52 \pm 7\%$ vs $25 \pm 8\%$ reduction; $p = .0132$), whereas LOW participants had an intermediate response (see Figure 2). The ES for the change in GDS outcome were 1.84, 1.37, and 0.75 for the HIGH, LOW, and GP care groups, respectively. Similarly, ES for the change in HRSD were 1.85, 1.21, and 0.70 for HIGH, LOW, and GP care groups, respectively. Using the standard pharmaceutical trial definition of a 50% reduction in HRSD to indicate a clinical response, HIGH was again approximately twice as effective (61% of participants with a clinical response) as either LOW (29% of participants; chi-square 3.728, $p = .05$) or GP care (21% of participants; chi-square

Table 2. Baseline Psychological Status

Variable	HIGH (N = 20)	LOW (N = 20)	GP Care (N = 20)	p Value
DSM-IV Diagnosis, n				.14
Minor depression	2	2	7	
Major depression	17	16	13	
Dysthymia	1	2	0	
Estimated duration of DSM-IV diagnosis, mo	29 ± 41	27 ± 31	19 ± 20	.4
Previous treatment with an antidepressant, n				.4
Yes	8	10	10	
No	12	10	10	
Precipitating cause of episode of depression, n				.9
Personal health	7	5	6	
Family illness	2	2	3	
Relationship difficulties	5	6	5	
Retirement	2	1	3	
Financial worries	2	2	1	
Bereavement	2	3	2	
None reported	0	1	0	
Social network, n				.5
Local family dependent	4	1	3	
Locally integrated	4	6	4	
Local self contained	1	1	4	
Wider community focused	2	5	4	
Private restricted	9	7	5	

Note: HIGH = high intensity progressive resistance training group; LOW = low intensity progressive resistance training group; GP = general practitioner; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders.

5.442, $p < .02$). LOW and GP care groups were not significantly different from each other (chi-square 0.332, $p = .56$) in this analysis of the proportion of participants with a clinical response. Thus, all three standard analyses of depression response (absolute scores, percent change, and $\geq 50\%$ reduction in HRSD analysis) support the hypothesis of a dose-response effect for PRT favoring higher intensity of exercise, and larger ES for both kinds of exercise relative

Table 3. Adverse Events

Variable	HIGH (N = 18)	LOW (N = 17)	GP Care (N = 19)	p Value
Visits to a health professional over the study (no. per person)	2.0 ± 2.0	2.0 ± 1.8	5.0 ± 1.8	<.0001
Minor illness (no. of weeks reported per person)	0.2 ± 0.4	0.2 ± 0.3	0.3 ± 0.3	.32
Muscular pain (no. of weeks reported per person)	4.1 ± 2.7	2.9 ± 2.6	3.6 ± 2.5	.34
Chest pain (no. of weeks reported per person)	0.9 ± 1.9	0.5 ± 0.9	0.5 ± 0.8	.4
Injuries requiring training adjustment (no. per person)	0.1 ± 0.3	0.2 ± 0.7	0	.9
Falls (no. per person)	0.15 ± 0.37	0.28 ± 0.75	0 ± 0	.2
Deaths	0	0	0	—
Hospital days (no. per person)	0.5 ± 0.2	0 ± 0	0.4 ± 0.5	.2

Note: HIGH = high intensity progressive resistance training group; LOW = low intensity progressive resistance training group; GP = general practitioner.

Table 4. Depression Outcomes

Variable	HIGH (N = 18)	LOW (N = 17)	GP Care (N = 19)	Time F, p Value	Group × Time F, p Value
GDS pre	20.0 ± 4.1	22.0 ± 4.3	18.7 ± 3.5	92.20, <.0001	5.58, .006
GDS post	8.4 ± 7.0	13.3 ± 7.0	14 ± 5.2		
HRSD pre	18.0 ± 4.5	19.5 ± 5.3	19.7 ± 3.9	74.87, <.0001	1.98, .14
HRSD post	8.5 ± 5.5	12.4 ± 6.3	14.4 ± 6.0		

Note: HIGH = high intensity progressive resistance training group; LOW = low intensity progressive resistance training group; GP = general practitioner; GDS = Geriatric Depression Scale; HRSD = Hamilton Rating Scale of Depression. F ratio and p values from repeated measures analysis of variance model.

to GP care. In the most rigorous analysis (response defined as a $\geq 50\%$ reduction in HRSD), low intensity exercise was neither statistically nor clinically different from GP care.

Secondary Outcomes

Muscle strength.—Mean strength gains (averaged across all muscle groups) were $37 \pm 3\%$ in HIGH, $6 \pm 1\%$ in

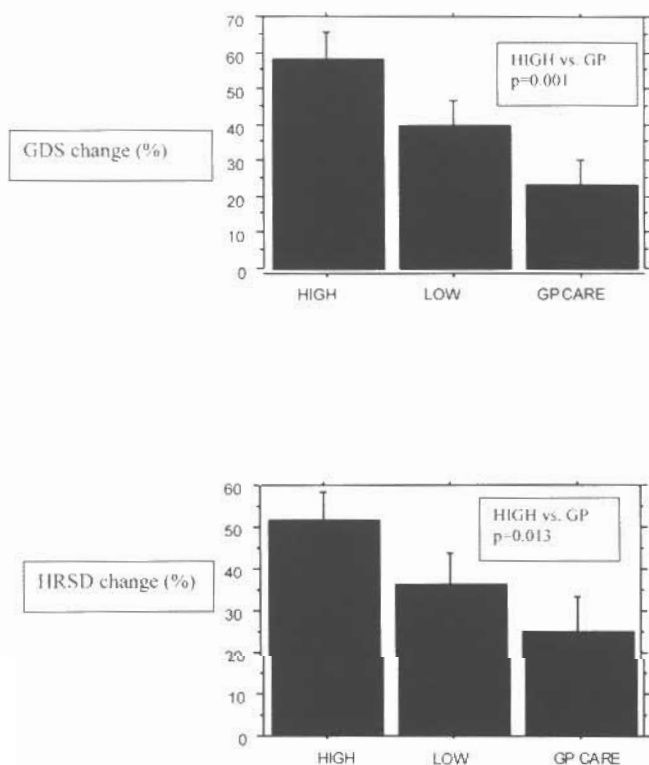


Figure 2. Relative change in self-rated and therapist-rated depression scales in the three study arms. Results were analyzed by analysis of variance (ANOVA) and post hoc *t* tests. **Top:** Reduction in self-rated depression. GDS = Geriatric Depression Scale; HIGH = high intensity progressive resistance training; LOW = low intensity progressive resistance training; GP = general practitioner standard care. $p = .004$ (ANOVA for group effect); $p = .001$ (HIGH vs GP; Fisher's protected least significant difference post hoc *t* test). **Bottom:** Reduction in therapist-rated depression. HRSD = Hamilton Rating Scale of Depression; HIGH = high intensity progressive resistance training; LOW = low intensity progressive resistance training; GP = general practitioner standard care. $p = .044$ (ANOVA for group effect); $p = .013$ (HIGH vs GP; Fisher's PLSD post hoc *t* test).

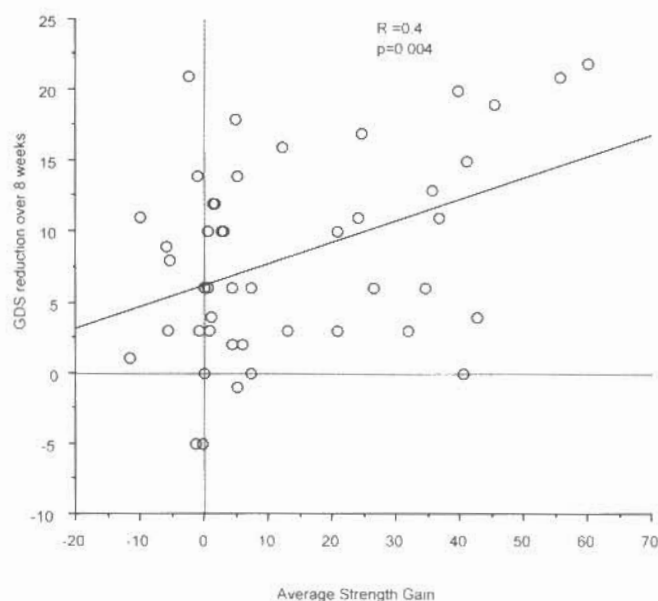


Figure 3. Relationship between change in depressive symptoms and gain in muscle strength. GDS = Geriatric Depression Scale. Average strength gain = mean of percent changes in six different resistance training exercises at 8 weeks.

LOW, and $-2 \pm 1\%$ in GP care groups ($p < .0001$), with all groups significantly different from each other in post hoc *t* test comparisons (HIGH vs LOW, $p < .0001$; HIGH vs GP, $p < .0001$; LOW vs GP, $p = .0036$).

Quality of life.—Six of the eight domains of the health-related quality of life questionnaire improved significantly across all groups after the intervention (p values ranging from $<.0001$ to $<.04$): Physical Function, Role Physical, Vitality, Social Function, Role Emotional, and Mental Health. In addition, there was a significantly greater effect ($p = .048$) of HIGH on Vitality (score 32.5 ± 17.0 to 61.4 ± 22.0), compared to changes in LOW (31.8 ± 14.6 to 46.8 ± 27.7) or GP care groups (29.2 ± 18.5 to 38.9 ± 17.7). This same trend for the greatest improvements to occur in HIGH was seen in all the other domains that improved as well (data not shown).

Sleep quality.—Scores on the Pittsburgh Sleep Quality Index and on five of its seven subscales improved significantly over time across all groups ($p < .0001$ for total score, $p < .0001$ for sleep-related quality of life, $p = .0166$ for sleep onset latency, $p = .0274$ for sleep duration, $p < .0001$ for sleep efficiency, and $p < .0001$ for sleep dysfunction). There was a significant effect of group assignment favoring HIGH over LOW and GP care on the percent change in overall sleep quality adjusted for baseline sleep quality ($p = .05$). The magnitude of this change in overall sleep quality was $24 \pm 8\%$ in the HIGH exercise group versus $22 \pm 7\%$ in LOW and $19 \pm 8\%$ in GP care.

Potential Predictors of Antidepressant Response

There was a significant relationship between strength gain and reduction in GDS in all groups combined (see Figure 3).

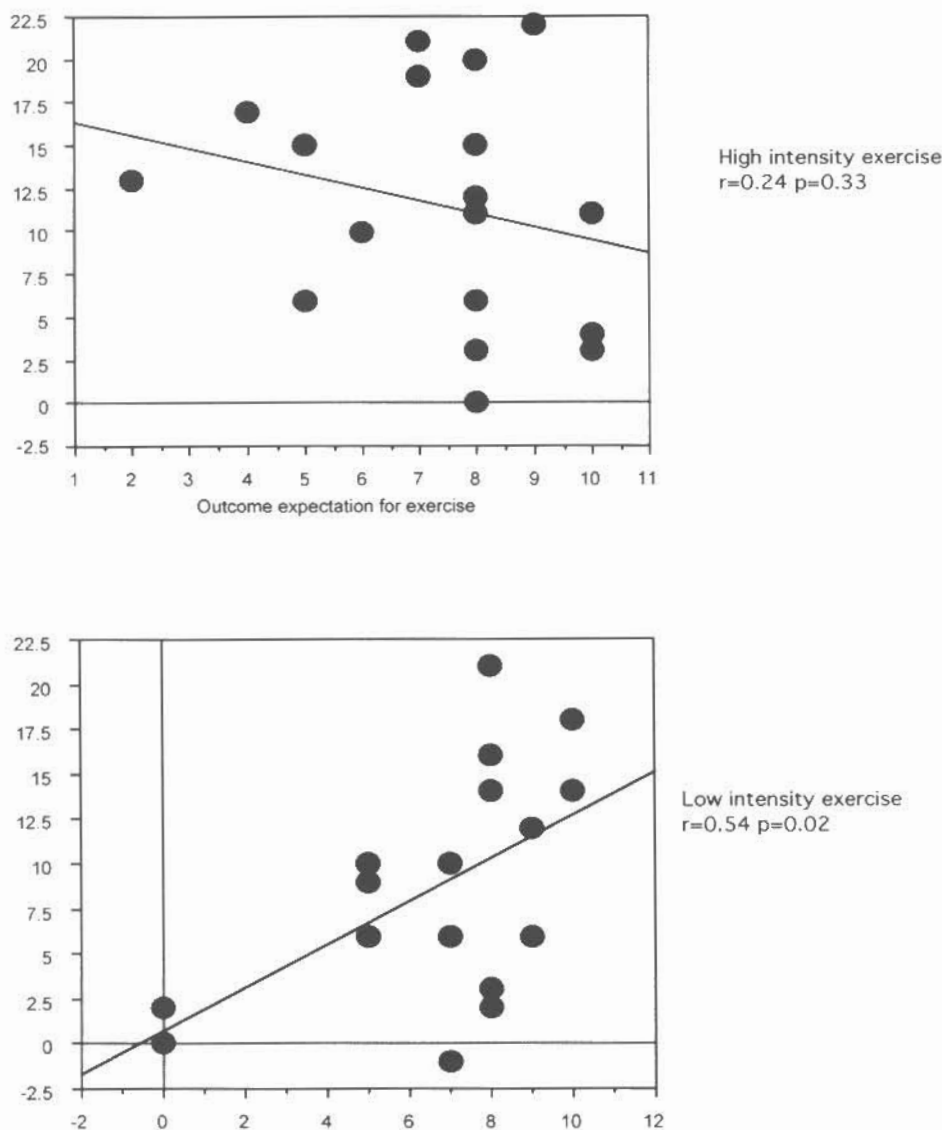


Figure 4. Relationship between outcome expectation and reduction in self-rated depression. **Top:** Results in high intensity progressive resistance training group. GDS 8-0 = absolute change in Geriatric Depression Scale score between baseline and 8 weeks. **Bottom:** Results in low intensity progressive resistance training group. GDS 8-0 = absolute change in Geriatric Depression Scale score between baseline and 8 weeks.

When analyzed separately by groups, this relationship was significant in HIGH ($r=0.53$, $p=.03$), but not in LOW ($r=.16$, $p=.56$) or GP care groups ($r=.001$, $p=.99$).

Baseline outcome expectation of GPs' ability to improve depression and participants' belief in efficacy of exercise were not significantly different between groups, and GP expectation did not predict response to any intervention. Outcome expectation for exercise efficacy had no relationship to reduction in depression in the HIGH group, but was significantly related to reduction of depression in the LOW group ($r=.55$; $p=.02$) (see Figure 4).

There was a significant effect of baseline social support network on the reduction in depression as assessed by the Hamilton Rating Scale of Depression ($F=3.52$, $p=.0152$), but there was no interaction between social network and treatment group effects ($F=.751$, $p=.631$). Those par-

ticipants classified as local family-dependent were the least responsive to any of the interventions, being significantly different than private restricted ($p=.045$), local self-contained ($p=.035$), and locally integrated ($p=.0024$) individuals. Social support network category did not change over time (data not shown).

Within the Locus of control of health, Internality and Chance did not differ at baseline between groups, showed no significant change over time or group by time interaction, and were not related to reduction in depression. The Powerful Others locus of control was significantly higher in the GP care group at baseline (HIGH 20.7 ± 5.4 , LOW 18.5 ± 4.8 , GP care 24 ± 5.4 ; $p=.01$). However, there was no significant effect of time or group by time interaction on Powerful Others locus of control, nor any relationship with depression outcomes.

There was a significant improvement in self-efficacy over time ($p = .003$) but no effect of group assignment (HIGH $35.8.7 \pm 8.2$ to 39.9 ± 7.7 ; LOW 30.8 ± 10.8 to 35.2 ± 11.4 ; GP care 30.7 ± 8.4 to 32.8 ± 9.3 ; $p = .64$). Self-efficacy (at baseline or change scores) did not predict antidepressant response. There were no associations between personality type at baseline, or between quality of life (baseline or change scores) and reduction in depression.

DISCUSSION

This is the first study to select clinically depressed elderly persons and demonstrate the efficacy and dose-response characteristics of weight lifting exercise for depression. The results of this trial suggest that high intensity weight lifting is an effective, feasible, and safe treatment for older depressed patients. There appears to be a dose-response effect of exercise on depression reduction, with high intensity being required for a clinically meaningful response. High intensity PRT was more effective than low intensity PRT (relative ES 0.64) and GP care (relative ES 1.15). There was no statistically significant difference in efficacy between low intensity exercise and standard care, with an average response of 21%–29%. This approximates the 30% response rate seen with placebo medications in acute episodes of major depression (5). It could be argued that standard GP care was suboptimal, with only 52% of patients receiving either pharmacotherapy or counseling, but this is similar to standard care in the community, according to the most recent U.S. data from the National Comorbidity Survey Replication report (2), in which only 52% of cases received treatment, with only 22% of cases classified as adequately treated. The GP care group had an average of five visits with health care professionals during the trial (more than twice that of the exercise groups), suggesting significant contact regarding their depression.

There is no previously published literature on the dose response effect of weight lifting exercise and depression. There is conflicting evidence on a dose-response effect for aerobic exercise. Increased energy expenditure in aerobic exercise was associated with a decreased incidence of depression in the Harvard alumni (17% risk reduction at 1000–2499 kcal/wk vs 28% risk reduction at ≥ 2500 kcal/wk) (31). Conroy and colleagues (32) reported that 3 days per week of aerobic exercise was more beneficial than 1 day per week in hospitalized patients with depression. In contrast, two randomized trials which contrasted equal frequencies of aerobic exercise at either high (jogging) or moderate (walking) intensity found equivalent depression responses in young depressed or "neurotic" participants (33,34). Finally, although two studies (7,35) support a relationship between change in aerobic capacity and reduction in depression scores, these relationships are modest and not uniformly seen (10). This topic is the subject of ongoing trials (36), to identify which of the elements of the aerobic "dose" (e.g., intensity, frequency, duration, total energy expenditure, or physiological adaptation) are linked to antidepressant efficacy. In our study, both the prescribed intensity of resistive exercise (relative load lifted) as well as the physiological adaptation (muscle strength gained) were significantly related to the magnitude of the depression response.

The type of aerobic exercise which has been primarily effective for depression (running, cycling, or jogging at moderate to high intensity) may not be suitable for some of the older depressed patients seen in general practice who may have multiple comorbidities which may limit effective participation in aerobic exercise at these intensities. Thus, the combination of better efficacy and a similar safety profile compared to low intensity resistance exercise, and feasibility in a broader range of patients than moderate-high intensity aerobic exercise, suggests that this may be the preferred mode of exercise treatment for this patient group.

Compared to standard antidepressant treatment of elderly persons (4), high intensity resistance training appears to offer similar efficacy, with a clinically meaningful response seen in approximately 60% of patients in this study, as in our previous work (8,12). However, given the poor side effect profile of antidepressant pharmacotherapy in older adults (5), and the low compliance with such treatment (4), the risk/benefit ratio and acceptability may favor resistance exercise.

Our results provide some direction for future research on the mechanism of exercise effects on depression. Socialization or attention does not appear to be the mechanism by which exercise works, which has been a widely held belief in this field. Our two exercise groups performed exercises that differed only in intensity, yet only the high intensity group achieved clinically significant benefits (29).

Does exercise work because patients believe in it (placebo effect)? Belief in the ability of exercise to reduce depression was not predictive of efficacy at high intensity, but it was at low intensity in this study. This finding suggests that low intensity exercise works like a placebo (29% response rate), with increasing benefit for those who believe in it most strongly. High intensity exercise works equally well, whether the person believes in it or not, suggesting that a different physiological pathway may be activated by this mode of treatment.

The only baseline characteristic that predicted treatment response was social support network, with individuals reporting a local family-dependent network faring the worst. However, social network did not change the pattern of response to the interventions, as high intensity exercise was still the most effective treatment across all categories. We found no evidence that self-efficacy, personality type, locus of control, or changes in health-related quality of life or self-efficacy were related to reductions in depressive symptoms, suggesting that these are not central to the mechanism of benefit of high intensity exercise.

Taken together, our findings point towards a biological mechanism of response. Known effects of exercise on hormonal milieu, neurotransmitter levels, and sympathetic and parasympathetic nervous system activity, all of which may be involved in the biology of major depression (37), make this a plausible pathway for future investigation.

High intensity exercise had additional benefits beyond relief of depression, including more significant improvements in health-related quality of life, sleep quality, and muscle strength than did low intensity exercise or GP care. This spectrum of benefits is broader than those attributable to standard care of depression, and of potentially great rele-

vance to older adults with physical frailty or sarcopenia in addition to their psychiatric diagnosis.

Summary

High intensity resistance training is superior to low intensity resistance training or usual care by a GP in older community-dwelling adults with clinical depression. Future studies are needed to examine whether this treatment is a viable alternative when directly compared to optimal pharmacological treatment of major depression in older adults, and to define which patients are most likely to benefit in terms of lasting disease remission and prevention of recurrent episodes.

ACKNOWLEDGMENTS

We would like to acknowledge and thank the patients and general practitioners in the Central Sydney Area Health Service for their participation and dedication to this work.

Address correspondence to Nalin A. Singh, MBBS, Balmain Hospital, Booth St., Balmain, NSW, Australia. E-mail: singhn@email.cs.nsw.gov.au

REFERENCES

- Lebowitz B. Diagnosis and treatment of depression in late life. An overview of the NIH consensus statement. *Am J Geriatr Psychiatry*. 1996;4(Suppl 1):S3-6.
- Kessler R, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003;289:3095-3105.
- Hirschfeld RA, Keller M, Panico S, et al. The national depressive and manic-depressive association consensus statement on the undertreatment of depression. *JAMA*. 1997;277:333-340.
- Mittmann N, Herrmann N, Einarson T, et al. The efficacy, safety and tolerability of antidepressants in late life depression: a meta-analysis. *J Affect Disord*. 1997;46:191-217.
- Salzman C, Schneider L, Lebowitz B. Antidepressant treatment of very old patients. *Am J Geriatr Psychiatry*. 1993;1:21-29.
- McNeil K, LeBlanc E, Joyce M. The effect of exercise on depressive symptoms in the moderately depressed elderly. *Psychol Aging*. 1991;3:487-488.
- Blumenthal JA, Babyak MA, Moore KA, et al. Effects of exercise training on older patients with major depression. *Arch Intern Med*. 1999;159:2349-2356.
- Singh NA, Clements KM, Fiatarone MA. A randomized controlled trial of progressive resistance training in depressed elders. *J Gerontol Med Sci*. 1997;52A:M27-M35.
- Doyle EJ, Ossip-Klein DJ, Bowman ED, Osborn KM, McDougall-Wilson IB, Neimeyer RA. Running versus weight lifting in the treatment of depression. *J Consult Clin Psychol*. 1987;55:748-754.
- Martinsen EW, Hoffart A, Solberg O. Comparing aerobic and non-aerobic forms of exercise in the treatment of clinical depression: a randomized trial. *Comp Psychiatry*. 1989;30:324-331.
- Singh NA, Clements KM, Fiatarone MA. A randomized controlled trial of the effect of exercise on sleep. *Sleep*. 1997;20:95-101.
- Singh NA, Clements KM, Singh MA. The efficacy of exercise as a long-term antidepressant in elderly subjects: a randomized, controlled trial. *J Gerontol Med Sci*. 2001;56A:M497-504.
- Yesavage JA, Brink TL. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*. 1983;17:37-49.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Vol 323. Fourth Ed. Washington, D.C.: American Psychiatric Association; 1994.
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.
- McDonagh MJN, Davies CTM. Adaptive response of mammalian skeletal muscle to exercise with high loads. *Eur J Appl Physiol*. 1984;52:139-155.
- Borg G, Linderholme H. Exercise performance and perceived exertion in patients with coronary insufficiency, arterial hypertension and vasoregulatory asthenia. *Acta Med Scand*. 1970;187:17-26.
- Dunn A, Trivedi M, O'Neal H. Physical activity dose-response effects on outcomes of depression and anxiety. *Med Sci Sports Exerc*. 2001;33:5587-5597.
- Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychiatry*. 1967;6:278.
- Hughes JR. Psychological effects of habitual aerobic exercise: a critical review. *Prev Med*. 1984;13:66-78.
- Eysenck H, Eysenck S. *Manual of the Eysenck Personality Questionnaire*. London, UK: Hodder and Stoughton; 1975.
- Wenger G. A network typology: from theory to practice. *J Aging Studies*. 1991;5:147-162.
- Wenger G, Tucker I. Using network variation in practice: identification of support network type. *Health Soc Care Community*. 2002;10:28-35.
- Sherer M, Maddux J. The self-efficacy scale: construction and validation. *Psychol Rep*. 1982;51:663-671.
- Wallston B. Development and validation of the Health Locus of Control (HLC) scale. *J Consult Clin Psychol*. 1976;44:580-585.
- Ware J. *SF36 Health Survey Manual and Interpretation Guide*. Boston, Massachusetts: Nimrod Press; 1993;10:14-33.
- Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1988;28:193-213.
- Buysse DJ, Reynolds CF, Monk TH, et al. Quantification of subjective sleep quality in healthy elderly men and women using the Pittsburgh Sleep Quality Index (PSQI). *Sleep*. 1991;14:331-338.
- North TC, McCullagh P, Tran ZV. The effect of exercise on depression. *Exer Sport Sci Rev*. 1990;18:379-415.
- G*Power: a priori, post-hoc, and compromise power analyses for the Macintosh [computer program]. Version 2.1.2. Trier, Germany: University of Trier; 1997.
- Paffenbarger R, Lee IM, Leung R. Physical activity and personal characteristics associated with depression and suicide in American college men. *Acta Psychiatr Scand*. 1994;377:16-22.
- Conroy R, Smith K, Felthous A. The value of exercise on a psychiatric hospital unit. *Hosp Community Psychiatry*. 1982;33:641-645.
- Sexton H, Maere A, Dahl N. Exercise intensity and reduction of neurotic symptoms: a controlled follow-up study. *Acta Psychiatr Scand*. 1989;80:231-235.
- Veale D, Le Fevre K, Pantelis C, De Souza V, Mann A, Sargeant A. Aerobic exercise in the adjunctive treatment of depression: a randomized controlled trial. *J Royal Soc Med*. 1992;85:541-544.
- Martinsen EW, Medhus A, Sandvik L. Effects of aerobic exercise on depression: a controlled study. *Br Med J*. 1985;291:109.
- Dunn A. The DOSE study: a clinical trial to examine efficacy and dose response of exercise as treatment for depression. *Control Clin Trials*. 2002;23:584-603.
- Singh NA. Depression in the older woman. In: Press C, ed. *Exercise, Nutrition, and the Older Woman. Wellness for Women Over Fifty*. Vol 1. Boca Raton, Florida: CRC Press; 2000:395-416.

Received December 17, 2003

Accepted March 17, 2004

Decision Editor: John E. Morley, MB, BCH