Mental Health and Physical Activity 6 (2013) 119–131

Contents lists available at SciVerse ScienceDirect

Mental Health and Physical Activity

journal homepage: www.elsevier.com/locate/menpa

Depressed Adolescents Treated with Exercise (DATE): A pilot randomized controlled trial to test feasibility and establish preliminary effect sizes

Carroll W. Hughes [a,*], Shauna Barnes [a], Conrad Barnes [a], Laura F. DeFina [b], Paul Nakonezny [a,c], Graham J. Emslie [a]

[a] Department of Psychiatry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-9044, USA
[b] Cooper Institute, Dallas, TX, USA
[c] Department of Clinical Sciences, Division of Biostatistics, University of Texas Southwestern Medical Center, Dallas, TX, USA

ARTICLE INFO

Article history:
Received 18 April 2012
Received in revised form
12 June 2013
Accepted 12 June 2013

Keywords:
Depression
Exercise
Adolescents
Adherence
Actigraphy

ABSTRACT

The Depressed Adolescents Treated with Exercise (DATE) study evaluated a standardized aerobic exercise protocol to treat nonmedicated adolescents that met DSM-IV-TR criteria for major depressive disorder. From an initial screen of 90 individuals, 30 adolescents aged 12–18 years were randomized to either vigorous exercise (EXER) (>12 kg/kcal/week [KKW]) or a control stretching (STRETCH) activity (<4 KKW) for 12 weeks. The primary outcome measure was the blinded clinician rating of the Children's Depression Rating Scale – Revised (CDRS-R) to assess depression severity and Actical (KKW) accelerometry 24hr/7days a week to assess energy expenditure and adherence. Follow-up evaluations occurred at weeks 26 and 52. The EXER group averaged 77% adherence and the STRETCH group 81% for meeting weekly target goals for the 12 week intervention based on weekly sessions completed and meeting KKW requirements. There was a significant increase in overall weekly KKW expenditures ($p < .001$) for both groups with the EXER group doubling the STRETCH group in weekly energy expenditure. Depressive symptoms were significantly reduced from baseline for both groups with the EXER group improving more rapidly than STRETCH after six weeks ($p < .016$) and nine weeks ($p < .001$). Both groups continued to improve such that there were no group differences after 12 weeks ($p = .07$). By week 12, the exercise group had a 100% response rate (86% remission), whereas the stretch group response rate was 67% (50% remission) ($p = .02$). Both groups had improvements in multiple areas of psychosocial functioning related to school and relationships with parents and peers. Anthropometry reflected decreased waist, hip and thigh measurements ($p = .02$), more so for females than males ($p = .05$), but there were no weight changes for either gender. The EXER group sustained 100% remission at week 26 and 52. The STRETCH group had 80% response and 70% remission rates at week 26 and by week 52 only one had not fully responded. The study provides support for the use of exercise as a non-medication intervention for adolescents with major depressive disorders when good adherence and energy expenditure (KKW) are achieved.

1. Introduction

1.1. Adolescent depression and treatments

Major depressive disorder (MDD) is common in adolescents and results in significant morbidity and mortality. In fact, adolescent suicide is the third leading cause of death for this age group and most are accounted for by depression (Hughes et al., 2007). With recent concerns about short and long term safety with SSRIs and other antidepressants, and poor overall response even when efficacy has been proven (Kratochvil, Vitiello, & Walkup, 2006), it is essential to determine whether reduction in depressive symptoms can be accomplished with a non-medication treatment. Clearly, any improvement in the treatment of depressed adolescents has important public health implications (Nemeroff, Kalali, & Keller, 2007) given the increasing prevalence of depression in childhood and adolescence and the relationship to obesity and its profound impact on functioning and long-term prognosis (Olfson, Marcus & Shaffer, 2006). In the past decade, selective serotonin reuptake inhibitor (SSRI) medications have received the most study in the

* Corresponding author. Tel.: +1 214 648 5277.
E-mail address: Carroll.hughes@UTSouthwestern.edu (C.W. Hughes).

1755-2966/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.mhpa.2013.06.006

treatment of depressed children and adolescents. There is an established history of successfully treating adolescents and children particularly with fluoxetine (Emslie et al., 2002, 2008; Emslie, Rush, Weinberg, Gullion, et al., 1997; Emslie, Rush, Weinberg, Kowatch, et al., 1997; Hughes et al., 2007), and fluoxetine now has FDA approval for use in children and adolescents. One consistent finding across the antidepressant studies of youth and adults is that around 30–40% do not respond to antidepressants (Birmaher & Brent, 1998; Brent, Kolko, Birmaher, Baugher, & Bridge, 1999). Many who do respond to treatment are left with residual symptoms of the illness (Kennard et al., 2006) such as poor quality of sleep and increased weight. Patients with residual symptoms have a more rapid onset of relapse or recurrence (Curry et al., 2011; Fava, Fabbri, & Sonino, 2002; Judd et al., 2000) and are more likely to seek and use health care services (Lin et al., 1998). Teens with more symptoms remaining at the end of 12 weeks of treatment are less likely to be remitted at 18 and 36 weeks post initiation of treatment and are more likely to relapse. Importantly, faster rate of response in the acute phase of treatment is associated with more complete improvement, with those responding to treatment within the first four weeks being more likely to achieve remission after 12 weeks of treatment (Brent et al., 2008; Tao et al., 2009). The major concern for the findings for both depressed adults and juveniles is that the medications are not without risk, remain a major health cost, and their use have declined due to warnings and side effects (Kratochvil et al., 2006), with a resultant increase in suicides (Nemeroff et al., 2007; Olfson et al., 2006).

1.2. Exercise as a treatment for adolescent depression

This study focuses on exercise as an important viable non-medication alternative to medication or psychotherapy (Oeland, Laessoe, Olesen, & Munk-Jergensen, 2010; Taliaferro, Rienzo, Pigg, Miller, & Dodd, 2009). It may be particularly attractive for a number of reasons. Some adolescents may be concerned with the stigma of psychotherapy or the side effects of medication. Additionally, evidence-based cognitive behavioral therapy (CBT) by certified therapists is not available everywhere and adolescents may not be able to reliably access such support. By contrast, the exercise intervention can be personalized to be compatible with each person's schedule, interests and fitness levels and can be done most anywhere without side effects, and continued as a life-long habit that may have numerous long lasting physical and mental health benefits (Dunn et al., 1999; Paluska & Schwenk, 2000; Penedo & Dahn, 2005).

A variety of evidence supports the use of exercise as an alternative to antidepressant treatment (Dunn, Trivedi, Kampert, Clark, & Chambliss, 2005; Rothon et al., 2010; Sund, Larsson, & Wichstrom, 2011; Trivedi et al., 2006). Exercise has been used both alone and in combination with other treatments for adult depression to alleviate depressive symptoms (Babyak et al., 2000; Blumenthal et al., 2007; Dunn et al., 2005), but such randomized controlled studies do not exist for depressed adolescents that meet criteria for DSM-IV-TR major depressive disorder. More recent meta-analyses of exercise studies and depression also support exercise as a treatment (Dunn & Weintraub, 2008; Lawlor & Hopker, 2001; Mead et al., 2009; O'Neal, Dunn, & Martinsen, 2000; Paluska & Schwenk, 2000). The most recent review for exercise in treating adult depressions suggests that findings from recent adult studies need to be interpreted with caution since the more methodologically rigorous studies have had smaller effects (Rimer et al., 2012). Multiple meta-analytic reviews of exercise studies for depression in young people have been used to argue that the small numbers of well-designed studies to date, and methodological concerns, limit any conclusions about the effectiveness of treating depressed youth with exercise in contrast with adult studies (Biddle, 2001; Larun, Nordheim, Ekeland, Hagan, & Heian, 2006).

Neuroscientific evidence provides plausible biological mechanisms by which exercise improves mood and reduces symptoms of depression (Dunn & Dishman, 1991; Dwivedi, Mondal, Rizavi, & Conley, 2005; Keller et al., 2010; Nabkasorn et al., 2006; Zarrilli et al., 2009). Recent human and animal studies have postulated that reduced levels of brain-derived neurotrophic factor BDNF play a major role in the pathophysiology of depression (Huang, Lee, & Liu, 2007), and that its restoration may represent a critical mechanism underlying antidepressant efficacy and/or improvement in depressive symptoms (Aydemir et al., 2006; Gonul et al., 2005; Karege et al., 2002). Physical activity has been associated with increased levels of neurotrophic factors, neuromodulators, and neuroamines, including brain-derived neurotrophic factors (Ferris, Williams, & Shen, 2007; Russo-Neustadt, Beard, & Cotman, 1999), noradrenergic and serotonergic systems (Meeusen, Piacentini, & DeMeirleir, 2001), and others (Szabo, Billett, & Turner, 2001), all of which have been extensively linked to depression (Russo-Neustadt et al., 1999; Siuciak, Lewis, Wiegand, & Lindsay, 1997).

1.3. Aims of the study

The aim of this study was to use a randomized control trial to test the primary hypotheses that 1) an aerobic exercise group (EXER) would have a quicker response with a lower depression severity (as measured by the CDRS-R total score) by week 6, 9 and 12 than a group that only did stretching exercise (STRETCH), and 2) that the EXER group would expend more energy weekly than those in STRETCH (as measured by Actical). It had been argued in the adult DOSE study (Dunn, Trivedi, Kampert, Clark, & Chambliss, 2002) that the stretching control group was serving as a placebo condition with only minimal energy expenditure. The DOSE stretch intervention had not been as effective as the EXER group in reducing depression severity and served as an important control for contact time with clinical and exercise staff. It was not deemed ethical to use a "wait-list" control with adolescents that met DSM-IV-TR criteria for a major depressive mood disorder.

The secondary hypotheses tested if the EXER group had greater improvements in various measures of depression self-report and psychosocial functioning by week 12 than STRETCH. A secondary aim was to establish effect sizes for the Chidren's Depression Rating Scale – Revised (CDRS-R) and Actical (energy expenditure data) as well as selected secondary outcomes (Ferguson, 2009). Further details of procedures and instruments have been published previously (Hughes et al., 2009).

2. Study procedures

2.1. Participant recruitment

Participants were recruited from the routine intake assessments of various Dallas, Texas outpatient clinics, and referrals from community clinics, physicians, and various website listings. Potential participants were screened by a phone interview which included a review of inclusion and exclusion criteria. At the enrollment visit, the principal investigator or a designee obtained IRB approved consent/assent for research participation from the potential subject's parent or legal guardians as well as the participant. A licensed study clinician then conducted separate semi-structured psychiatric diagnostic interviews schedule for Affective Disorders and Schizophrenia for school-age children – present and life-time [KSADS-PL]; (Kaufman et al., 1997) first with the adolescent and then the parent to confirm the diagnosis and current severity and family psychiatric history, reviewed all inclusion and exclusion

criteria, and assured that all of the baseline self-report instruments were completed. The case was then presented at a weekly research consensus conference of the team of child psychiatry clinical researchers to confirm diagnoses.

2.2. Inclusion criteria

The adolescent had to be between the ages of 12–18 years, in school and meet a DSM-IV-TR diagnosis of non-psychotic major depressive disorder (MDD). Both genders were included. While MDD was the primary disorder, youth with other concurrent disorders (anxiety, attention deficit [ADHD], or conduct) were not excluded. Children with ADHD and on medication were included if clinically stable on existing nonantidepressant medications. Participants needed a baseline score on the Children's Depression Rating Scale - Revised (CDRS-R) ^{3} 35 $\leq$ 70; and a baseline Clinical Global Impression – Severity (CGI-S) ^{3} 4. They had to be of normal intelligence (i.e., IQ $>$ 70 based on the short WISC-III testing if uncertain) and English language and reading were required as the majority of the assessment instruments were only available in English. In addition they had to provide a letter from a family physician approving that it was safe to participate in an exercise study. The exercise physician (LFD) at the Cooper Institute (CI) site in Dallas, Texas reviewed the information and made the final approval for the participant to exercise consistent with the American College of Sports Medicine [ACSM] guidelines.

2.2.1. Anthropometry

Measures of body mass, stature, and distribution of subcutaneous fat were taken at baseline and week 12 by CI staff. Waist and hip circumferences (i.e., waist/hip ratio) were assessed for determining body fat distribution, as was skin fold fat determined at two trunk sites (abdominal, suprailliac) and at two extremity sites (triceps, front thigh) using standardized procedures.

2.2.2. Randomization

Following final approval, all participants who had met inclusion criteria for the study were stratified based on gender and ADHD diagnosis (yes/no) and then randomized in a 1:1 ratio into one of the two possible treatment groups (EXER or STRETCH). The randomization schedule was produced using a random number program (SAS software via PROC PLAN) by a co-author (PN).

2.3. Exclusion criteria

Individuals were excluded if they had a chronic medical illness requiring regular medications that contraindicated intensive exercise or unstable medical conditions requiring medication(s) with psychotropic effects (e.g., anticonvulsants, steroids). Individuals with psychiatric diagnoses of lifetime history of any psychotic disorder, including psychotic depression, bipolar I and II disorder, schizophrenia, alcohol or substance abuse or dependence within the past six months, lifetime anorexia nervosa or bulimia were excluded. Anyone with severe suicidal ideation or previous history of serious suicide attempts, or severe depression (CDRS $>$ 70) that clinically needed immediate medical intervention were excluded as were those who participated in current exercise defined as 30 min of vigorous physical activity, 5 times per week or more based on the initial phone screen with the participant.

2.4. Measures

The primary outcome measure of depression severity was based on the blinded clinician rating of the CDRS-R total score at baseline and weeks 3, 6, 9, and 12 and follow-up at weeks 26 and 52. The weekly total amount of energy expended was based on the 24/7 ACTICAL monitor data (see below). The Actical data record was reviewed and recorded weekly with the participant to assess energy expenditure and adherence with required sessions of exercise. The other measures (psychosocial functioning and exercise evaluation measures) were completed at baseline, week 6, and after the study was completed at week 12 or at early exit.

2.5. Childhood Depression Rating Scale – Revised [CDRS-R]

The Children's Depression Rating Scale - Revised [CDRS-R] is a clinician-rated instrument, modeled after the Hamilton Depression Rating Scale for adults, and used to measure the presence and severity of depressive symptomatology in children and adolescents. The 17 items of the scale are rated on a 1 to 5 or 1 to 7 point scale, with a 1 describing absence of the given symptom. The CDRS-R yields a total score from 17 to 113 with a score of 36 or greater considered being compatible with a diagnosis of depression. Good test-retest reliability (.80), internal consistency (Cronbach Alpha = .85) and interrater reliability (.92) have been demonstrated for the instrument (Poznanski & Mokros, 1996).

2.6. Clinical Global Impression scale (improvement [CGI-I] and severity [CGI-S])

The Clinical Global Impression scale (CGI) was administered blind and was used to assess overall clinical severity (CGI-S) and improvement (CGI-I), each with a seven point scale, with lower values being more favorable. At intake, only severity can be rated. In subsequent assessments, both severity and improvement were rated (Guy, 1976). This is a standard scale for psychopharmacological research, and a CGI-Improvement of 1 (very much), or 2 (much) improved, is considered to be an acceptable response to acute treatment as is a clinical severity rating of less than or equal to 3.

2.7. Children's Global Assessment Scale (C-GAS) and Family Global Assessment Scale (F-GAS)

The Children's Global Assessment Scale (C-GAS) was administered blind and was adapted from the Global Assessment Scale for Adults (Shaffer et al., 1983). The C-GAS is included since it provides a measure of the overall level of functioning not limited to impairment from depression. The Child's Family Global Assessment Scale (F-GAS) rates the child's family's most impaired level of general functioning in the past year. There are four areas of functioning to be considered when rating: Social functioning of parents as related to economic and social goals; marital/parental teamwork; parent understanding and provision for the developmental needs of the child; integrity and stability of family relationships.

2.8. Quick Inventory for Depressive Symptomatology – adolescent version; clinician-rated & self-report (17 item; QIDS-A-C_{17}, QIDS-A-P_{17} and QIDS-A-SR_{17})

The Quick Inventory of Depressive Symptoms – Adolescent version (Clinician report, self-report and parent version) consists of 17 items on each instrument designed to assess both the core criterion symptoms and the commonly associated symptoms of depression (Rush, Bernstein, & Trivedi, 2006). The QIDS correlates highly with both the HAM-D and the Beck Depression Inventory (BDI) (Rush, Carmody, & Ibrahim, 2006). The blinded QIDS-A_{17}-C was included to allow comparison with adult studies, and because it contains items relating to atypical depression, such as reactivity of mood, rejection sensitivity, hypersomnia, and weight gain, which

may be more common in adolescents and may potentially change with exercise augmentation (Bernstein et al., 2010).

Psychosocial and exercise measures e.g., Profile of Mood State [POMS; (Curran, Andrykowski, & Studts, 1995)], Physical Activity Enjoyment Scale [PACES; (Kendzierski & DeCarlo, 1991)], and Social Adjustment Scale-Self-Report [SAS-SR; (Weissman & Bothwell, 1976)] are covered in more detail in a prior report (Hughes et al., 2009).

2.9. Physical activity – actical kilocalories (kcal)

Total energy expenditure for each exercise session and the total activity of each week of the study was reviewed with the participant at each weekly session (any inactive period that reflected that the device was not being worn was addressed with a re-emphasis of importance of wearing the device at all times) and used to calculate percent (%) adherence (e.g., kcal prescribed/kcal completed × 100). The Actical (Mini-Mitter, Bend, OR) monitored activity 24 h per day, seven days per week for the duration of the 12 week study for each participant. The Actical provides calculated energy expenditure values for Active Energy Expenditure in kilocalories/kilogram and total energy expenditure in Metabolic Equivalents per Time (METs) in kilocalories/min/kilogram. The device is a compact, battery-operated, physical activity monitor with physical characteristics similar to a small wristwatch that provided an objective measure of exercise adherence and total energy consumed.

A subsample of both exercise and stretch subjects were measured at the Cooper Institute while conducting their prescribed session with a Polar RS400 heart rate monitor and Actical. This allowed for comparison of the heart rates for the STRETCH versus EXER routines and estimates of their percent of highest average maximum heart rate expenditure (220 − age = maximum heart rate).

2.10. Timing of assessments

Clinical research staff closely monitored all study patients weekly for the duration of the study for any significant clinical worsening or concerns related to suicidality as part of the weekly outcome measures along with any adverse events related to the intervention. Ongoing weekly clinical assessment was sustained by the QIDS-A outcome measure collected at weekly intervals and the CDRS/CGI-I tri-weekly for the 12 weeks of exercise/stretching treatment.

All of the additional functioning measures were based on the repeated outcome measures at the end of the 12 weeks of EXER or STRETCH, or gathered at the time of early termination. Participants were also evaluated at 6 and 12 months. The adolescent participants received a cash equivalent of $25 upon completion of each weekly EXER or STRETCH visit at CI to compensate their time and travel expenses, resulting in a possible total payment of $300 if all 12 weekly assignments were completed. Those who returned for a 6 and 12 month follow-up evaluation received $50 at each visit.

2.11. Intervention (EXERcise or STRETCHing)

The amount of time required for participating in the exercise activities was the same for the EXER group and the STRETCH group. The only difference was the amount of energy expended during the activity. At the first session, the exercise trainer explained the procedures for the respective intervention (EXER or STRETCH), showed them the equipment available for the exercise or stretch sessions, and the coordinator familiarized the participant with the Actical device. The first two weeks required a minimum of 3 sessions at CI for the trainer to teach them how to use the equipment and complete the exercise or stretch routines. Following the first 2 weeks, participants began doing their exercise program at home or other location (gym, park, etc.), and only had to come to CI once a week for an exercise session. Each EXER/STRETCH session averaged about 30–40 min.

2.11.1. EXERcise intervention

Supervised exercise sessions at the Cooper Institute (CI) for the participants began by using the treadmills or stationary cycles. The CI trainers also taught patients how to complete home-based exercise sessions (e.g., choice of Wii Sports and Fit, jazzercise, jogging, weight training based on their preferred exercise) that were unsupervised workouts at the patient's home or in the community. The duration of each session generally was the time required to reach 1/3 or 1/4 of the total weekly caloric expenditure. There was a progression to the assigned exercise dose in the first few weeks that got them up to their minimum of 12 kilocalories/kilogram/week (KKW) energy expenditure (e.g., 8 KKW first week, 10 KKW second and 12 KKW by the third week). Participants exercised three times per week.

2.11.2. STRETCH intervention

The stretch group spent approximately the same amount of time, but at energy expenditures of less than 4 KKW per session. After two weeks of three sessions at CI they moved to once a week at CI and two home-based sessions. A 5–10 min stretching warm-up period included stretches that exercise the major muscle groups of the body. The series included such traditional "warm-up" stretches as: stretches of the gluts, inner thigh, calves and ankles, Achilles tendon, hamstring stretches, shoulder rolls forward and back, shoulder shrugs, isometrics for the neck hugging knees into the chest, moving forehead to right knee, then to left, then to both, and use of the pelvic tilt. An additional 10–15 min consisted of moving on to right and left calf stretches, quad stretches, and then to a series for the arms, hands, fingers, wrist, biceps/triceps, shoulders and back. All of the exercises were designed to be done slowly, emphasizing proper alignment, and rest periods to minimize overall physical exertion while obtaining general flexibility, and most importantly controlling for contact time with trainers and any social facilitation from participating in such activities. We had a different set of low level/low intensity routines for each of the 12 weeks to minimize boredom with the routines (Hughes et al., 2009).

2.12. Adherence to the exercise protocol

Exercise adherence was a concern in a study such as this one (Biddle, 2001; Dunn, Trivedi, & O'Neal, 2001). A major developmental difference in working with younger patients was the expectation that the parents would help encourage and monitor the regular exercise sessions. It was also critical that they endorse the importance of regular activity for their child and provide reinforcement for participating and completing non-CI exercise sessions. Staff regularly called and reminded participants of pending appointments and followed up on missed exercise sessions. CI had a well-designed and tested process for scheduling make-up sessions to sustain the goals of kcal expenditure for the study. Improved adherence was facilitated by developing and maintaining an excellent therapeutic alliance with participants and parents early on and by weekly monitoring of adherence to intervene with problem solving and support as needed. Given that it was unrealistic to expect 100% adherence to the exercise protocol in adolescents, acceptable adherence was defined as completing 70%

or more of the total amount of exercise prescribed for any given week as with previous adult studies (Dunn et al., 2002; Morss et al., 2004).

2.13. Documentation of treatment fidelity

Whether exercising at CI or home, participants were asked to keep an online diary and written exercise/stretch logs of frequency and duration of all exercise/stretch sessions. The DATE interactive website (see Hughes et al., 2009) was used to schedule and document each exercise session throughout the 12 weeks of the study. At the weekly CI visit the trainer and study coordinator went over the prior week's log and EXER/STRETCH sessions to review compliance and confirm that the weekly energy goal had been met, or in the STRETCH group, that the sessions had been completed. The CI Trainer provided feedback about how each participant was doing, reinforcement for goal attainment, and encouragement for those who needed to improve.

2.14. Client Satisfaction Questionnaire-8 (CSQ-8)

Subject satisfaction and acceptability of the interventions was based on a self-report exit interview. The CSQ-8 (Nguyen, Attkisson, & Stegner, 1983) has been used in other adolescent treatment studies demonstrating adequate internal consistency.

2.15. Data analyses

Baseline measures were made at the point the participant completed the initial assessment and met all criteria for study inclusion. After this, participants were then randomized to treatment. Baseline demographic and clinical characteristics for the sample are described using the sample mean and confidence intervals for continuous variables and the frequency and percentage for categorical variables. Two-independent samples *t*-test (for continuous variables) and Chi-square tests (for categorical variables) were used to compare the two exercise treatment groups (EXER vs. STRETCH) on the various demographics and clinical characteristics, including response and remission.

The change over time (12 weeks) in depressive symptoms (CDRS-R total), weekly energy expenditure, and psychosocial functioning was compared between the two exercise treatment groups (EXER vs. STRETCH) using a linear mixed model regression analysis of repeated measures. A separate model was conducted on each outcome measure. The model contained fixed effects terms for exercise treatment group, gender, time, and treatment group × time interaction. Baseline scores were included as covariates in each model. Where a measurement instrument had a joint set of rating subscales (e.g., PACES), a preliminary multivariate analysis of variance (MANOVA) was carried out to examine the effect of EXER vs. STRETCH at week 12 to minimize Type I error followed by appropriate post hoc tests.

All analyses were carried out using SPSS software, version 19 (IBM Corporation, Armonk, NY, USA). The level of significance for all tests was set at = .05 (two-tailed). SPSS calculates the effect sizes based on the partial η^2 where .01, .06, and .14 are considered small, medium, and large effects, respectively (Green & Salkind, 2011). One index of the relative impact of treatment conditions is effect sizes which can be useful in determining the clinical significance of findings and are critical in evaluating the stability of results across samples, designs, and analyses (Wilkinson, 1999). Reporting effect sizes also informs power analyses and meta-analyses needed in future research and was a secondary aim of this study since prior effect sizes for these measures have not been established for exercise in depressed youth.

3. Results

3.1. Study attrition

Of the 90 patients that received a phone screen evaluation, 40 were consented, and 30 were randomized. Of the 30 randomized, 26 completed the 12 week trial. Fig. 1 indicates the reasons for four youth not completing the study who had no, or very minimal initial data, and were therefore not included in the analyses. Twenty one of the individuals completed a 6 month follow-up evaluation and 15 completed a 12 month evaluation.

3.2. Participant characteristics

Between groups there was a mixture of gender (67% male, $p = .4$) and ethnic representation with 58% Caucasian, 25% Hispanic, and 17% African-American ($p = .7$). The adolescents had a mean age of 17 years (16.1–17.9; no group difference; $p = .9$) with a predominantly middle class (67%; $p = .5$) representation. Forty-two percent of the subjects had MDD only, 33% had MDD plus dysthymia, 33% had MDD plus ADHD, and one with MDD also had an anxiety disorder (Table 1). The majority of the participants were on no psychotropic medication, but two were entered on stable doses of fluoxetine since their depression had failed to fully respond and they continued to meet study entry criteria. Two were started on antidepressant medications (one in each group) six weeks after entering the study at the parents' initiative although improvement was seen by then and they were not exited from the study. Eight were entered who were on a stable dose of stimulant medication for their concomitant ADHD disorder (Table 1).

3.3. Effects of treatment on depression

The blinded clinician rated CDRS total score of depression severity was considered the major outcome measure for depression and was found to be significantly reduced over the 12 weeks for both groups based on a group by time interaction [$F(3,19) = 3.6$, $p = .034$, $\eta^2 = .36$] using a mixed model of linear regression repeated measure analysis that included baseline CDRS-R total scores as a covariate and group and gender as main effects (Fig. 2). The EXER group responded more quickly than the STRETCH group with lower depression scores based on a main effect for group [$F(1,21) = 5.5$, $p = .029$, $\eta^2 = .21$]. Males and female did not differ from each other by group or in terms of speed of initial response or in speed of response over the weeks of treatment. Post hoc tests found the biggest differences for groups at weeks 6 ($p < .016$) and 9 ($p < .001$), with the EXER group responding more rapidly in reduction of depressive severity but both groups improving from baseline. The groups did not differ by week 12. A similar finding was found for the blinded Clinician's QIDS and CGI – Improvement and Severity scores (see Table 2). The participant's QIDS self-report did not differ between groups but did show a significant reduction over time [$F(4, 21) = 12.3$, $p = .000$, $\eta^2 = .7$] as did the parent's [$F(4,21) = 10.4$, $p = .000$, $\eta^2 = .67$. There were no significant correlations for CDRS total scores with energy or activity counts.

3.3.1. Clinical response and remission

Equally important were the findings for clinical response and remission (see Table 3). Clinical response in prior depression studies (Emslie et al., 2008; Tao et al., 2009) has been defined as a CGI-I score less than equal to two and a CDRS total score with a greater than 50% reduction from baseline. The EXER group had a 100% response rate, whereas the STRETCH group's response rate was 67% ($\chi^2 = 5.52$, df = 1, $p = .019$). Remission on the other hand requires the CGI-I to be less than equal to two and the CDRS total

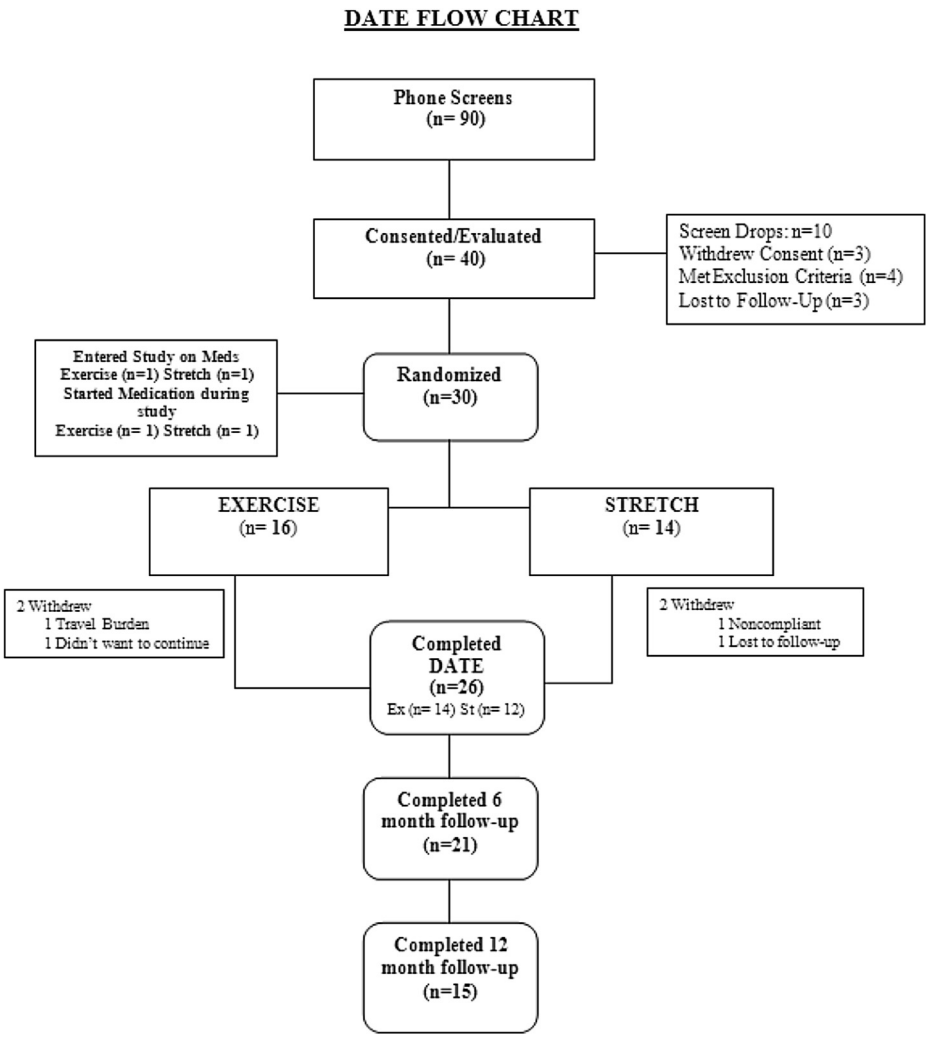


Fig. 1. Study enrollment and retention consort chart.

score being less than equal 28. The EXER group achieved an 86% remission rate compared to 50% for the STRETCH condition ($\chi^2 = 3.9$, df = 1, $p = .049$). Follow-up evaluations at weeks 26 and 52 indicated a sustained response-remission state for both groups. These rates exceed those typically found in antidepressant studies (Brent et al., 2008; Emslie et al., 2002; Emslie, Rush, Weingberg, Kowatch, et al., 1997; March et al., 2004).

3.3.2. *Profile of Mood State – Short (POMS)*

There are five factors that have been derived from the 27 items of the POMS. Each was analyzed separately: depression, anger, fatigue, tension, and vigor. There were no group differences for any of the factors; all except the Vigor factor showed significantly reduced scores by week 12: Depression – $F(2,23) = 114.1$, $p = .0001$, $\eta^2 = .55$; Anger – $F(2,23) = 4.1$, $p = .03$, $\eta^2 = .26$; Fatigue – $F(2,23) = 10.0$, $p = .001$, $\eta^2 = .47$; and Tension – $F(2,23) = 27.7$, $p = .0001$, $\eta^2 = .71$.

3.4. *Energy expenditure*

Total energy expenditure for each group can be seen in Fig. 3. The weekly total Kcal/Kg expenditure for each group clearly indicates significant increases from baseline level [M(C.I.) = 5959 (4781–7136) KKW] in energy expenditure, also supporting good adherence to the protocol (Fig. 3). The groups did not differ in baseline energy expenditure based on a week of Actical recording before beginning intervention. There was a group difference for the 12 weeks of intervention [$F(1,18) = 5.5$, $p = .03$, $\eta^2 = .23$], but no time or group by time interaction. The EXER group more than doubled the amount of energy expended, representing their 12 KKW dose as well as their activity counts (Fig. 4), and consistently so throughout the study. Interestingly there was a small increase in the STRETCH group energy expenditure from baseline but well within the less than 4 KKW goal.

3.4.1. *Heart rate monitoring*

Data from the Polar RS400 heart rate monitor indicated an average resting state heart rate of 74 (68.9–79.2). The average maximum heart rate for EXER was 165 (159–170) bpm or approximately 81% max, and for STRETCH 145 (138–152) bpm or 71% max which was significantly different from each other [$F(1,124) = 17.6$, $p < .001$, $\eta^2 = .12$].

3.5. *Treatment adherence*

Adherence was determined objectively by the total energy expenditure for prescribed weekly EXER sessions or the number of STRETCH sessions completed, and was verified at the weekly evaluation. These measures determined what percent adherence was

Table 1
Baseline sample demographics and clinical characteristics.

	Total ($n = 26$)	Exercise ($n = 14$)	Stretch ($n = 12$)
Mean age years (confidence interval)	17 (16.1–17.9)	17 (16.3–17.9)	17 (16.3–17.6)
Range	12–18	13–18	12–18
Male	16.7 (15.8–17.6)	16.6 (15.3–17.9)	16.9 (15.7–18.1)
Female	17.3 (16.1–18.3)	17.4 (16.1–18.7)	17.0 (15.3–18.7)
Gender, *n* (%)			
Male	15 (58%)	7 (50%)	8 (67%)
Female	11 (42%)	7 (50%)	4 (33%)
Ethnicity, *n* (%)			
Caucasian	15 (58%)	8 (57%)	7 (58%)
Hispanic	6 (23%)	2 (14%)	3 (25%)
African American	5 (19%)	4 (29%)	2 (17%)
Socio-economic-status, *n* (%)			
Low	1 (4%)	– (0%)	1 (8%)
Middle	19 (73%)	11 (79%)	8 (67%)
High	6 (23%)	3 (21%)	3 (25%)
Co-morbid diagnoses, *n* (%)			
MDD[a] only	14 (54%)	9 (64%)	5 (42%)
MDD + Dysthymia	5 (19%)	1 (7%)	4 (33%)
MDD + Anxiety Disorder	1 (4%)	– (0%)	1 (8%)
MDD + Behavior ADHD[b]	7 (27%)	3 (21%)	4 (33%)
Medications, *n* (%)			
None	18 (69%)	10 (71%)	8 (57%)
Antidepressants	2 (8%)	1 (4%) Citalopram	1 (4%) Fluoxetine
Stimulants	8 (31%)	3 (12%)	5 (19%)
Number who started post 6 weeks	2 (8%)	1 (4%) Zoloft	1 (4%) Fluoxetine
Number who started post 12 weeks	3 (12%)	2 (8%) Fluoxetine	1 (4%) Fluoxetine
Stimulants added post 12 weeks	3 (12%)	1 (4%)	2 (8%)

[a] DSM-IV-R major depressive disorder.
[b] Attention deficit hyperactivity disorder.

achieved (see Fig. 5). The EXER group averaged 10.0 out of 12 possible paid sessions (@ $25/session) and the STRETCH group 10.2 sessions (@ $25/session). The EXER group averaged 77% adherence and the STRETCH group 81% for the 12 weeks. Two consecutive weeks of less than 75% adherence to their prescribed weekly EXER dose or for number of STRETCH sessions completed was a basis for excluding the individual from the study. There was an overall tendency for adherence to trend down over time with an upswing at a major evaluation at the 6 week point. Nonetheless all participants, with the exception of one who was exited early on, managed to meet 12 weeks of adherence to the prescribed EXER or STRETCH regimen.

3.6. Effects of treatment on anthropometry measures

The summary of anthropometric findings summarized in Table 4 is informative. One of the female subjects in the EXER group was removed from these measures as her weight was four standard deviations out from the mean and instruments were not available to capture all of her other physical measurements. She did successfully complete and respond to the 12 weeks of the intervention.

Males were taller than the females (see Table 4 for the statistical values not reported in the text). Weight did not decrease and the males in both the EXER and STRETCH group gained an additional 4–5 pounds; some of the males had included weight training as part of their exercise. Although weight was minimally affected,

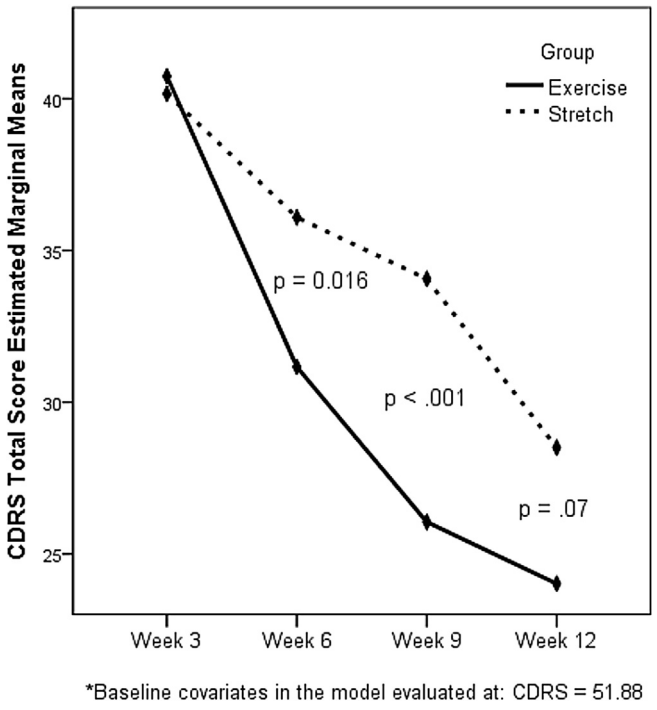


Fig. 2. The blinded clinician rated CDRS total score of depression severity by group.

significant reduction in waist circumference did occur in the females in the STRETCH condition (3 ½″) but not for exercise. The hips, thighs, and abdominal measurements decreased for females in the EXER condition but increased with STRETCH. Males had lower BMI and percent body fat than females with neither EXER nor STRETCH affecting these values by week 12. Females had lower bone density than males which did improve significantly following the EXER intervention.

3.7. Exercise measures

The Physical Activity Enjoyment Scale (PACES) rated attitudes and how the participants felt about various aspects of exercise [see (Hughes et al., 2009) for details]. A multivariate analysis of variance (MANOVA) was conducted for the 18 items measured at baseline and again at week 12. There were no significant differences for groups on any of the items at either time. A linear repeated measures analysis of variance did not find any differences between the baseline and week 12 scores. On a scale of 1–7 where 1 is favorable (e.g. enjoy exercise), 4 neutral, and 7 is negative (e.g. hate exercise), the overall mean for all 18 items at baseline for EXER was 3.75 (3.5–4.0) and for STRETCH 3.84 (3.6–4.1). By week 12 the EXER mean was 3.69 (3.5–3.9) and STRETCH 3.74 (3.5–4.0).

3.8. Psychosocial functioning

The adolescent Clinical Global Assessment of functioning (C-GAS) indicated equal improvement over time for both groups [baseline = 61.6 (59.5–66.7) to week 12 = 71.1 (55.8–63.7)], $F(2, 23) = 83$, $p < .001$, $\eta^2 = .88$. There was an interaction of time by group for the Family Global Assessment of function (F-GAS) with greater improvement for the EXER group [$F(8.5\ (2, 23) = 8.45$, $p = .002$, $\eta^2 = .61$).

3.8.1. Social Adjustment Scale – Self Report (SAS-SR)

Out of the 23 items on various aspects of social functioning, factors were created for school performance, relations with friends,

Table 2
Means (95% confidence intervals), and *p*-values for tri-weekly measures by treatment condition.

Measure	Exercise ($n = 14$)	Stretch ($n = 12$)	p values[a]
	Mean (95% CI)	Mean (95% CI)	
CDRS – clinician			
Week 0	50.9 (47.5–54.2)	53.6 (49.9–57.2)	.268
Week 3	40.0 (35.9–44.3)	41.3 (37.2–46.8)	.653
Week 6	30.8 (27.2–34.2)	37.1 (33.6–41.6)	**.016**
Week 9	26.0 (23.7–28.4)	33.1 (31.6–36.9)	**.001**
Week 12	24.1 (20.8–27.5)	28.3 (24.6–32.2)	.071
QIDS-clinician report			
Week 0	13.3 (11.8–14.8)	14.4 (12.8–16.0)	.304
Week 3	9.8 (8.4–11.2)	10.4 (8.9–11.9)	.526
Week 6	6.4 (5.3–7.6)	8.8 (7.5–9.9)	**.008**
Week 9	5.2 (3.8–6.6)	7.0 (5.5–8.5)	.08
Week 12	4.4 (2.7–6.0)	5.6 (3.8–7.4)	.305
QIDS-self-report			
Week 0	9.1 (6.1–12.0)	10.0 (6.8–13.2)	.662
Week 3	6.5 (4.4–8.6)	5.8 (3.6–8.1)	.657
Week 6	4.7 (2.7–6.7)	6.2 (4.0–8.3)	.662
Week 9	5.2 (2.8–7.7)	6.0 (3.4–8.6)	.657
Week 12	2.9 (1.2–4.5)	4.5 (2.7–6.3)	.174
QIDS-parent report			
Week 0	10.0 (8.0–12.0)	12.8 (10.6–14.9)	.066
Week 3	5.6 (3.3–8.0)	9.7 (7.1–12.2)	**.026**
Week 6	5.2 (3.3–7.2)	7.9 (5.8–10.0)	.064
Week 9	5.5 (3.3–7.7)	6.4 (4.0–8.8)	.330
Week 12	4.1 (2.4–5.9)	5.2 (3.3–7.1)	.420
CGI – depression severity			
Week 0	4.4 (4.1–4.6)	4.7 (4.4–5.0)	.125
Week 3	3.4 (3.0–3.8)	3.7 (3.2–4.1)	.402
Week 6	2.6 (2.3–2.9)	3.0 (2.7–3.3)	.062
Week 9	1.9 (1.6–2.3)	2.8 (2.4–3.1)	**.002**
Week 12	1.4 (1.0–1.9)	2.1 (1.6–2.5)	**.04**
CGI – improvement			
Week 0	–	–	
Week 3	3.1 (2.7–3.4)	3.0 (2.7–3.4)	.790
Week 6	2.4 (2.0–2.7)	2.7 (2.3–3.1)	.125
Week 9	1.9 (1.5–2.2)	2.3 (1.8–2.7)	.17
Week 12	1.2 (.9–1.6)	1.8 (1.4–2.3)	**.04**

[a] Independent *t*-tests, alpha uncorrected.

relations with family, dating, and general anxiety. All of these factors except dating showed improvement over the 12 weeks: SCHOOL – $F(1,24) = 4.6$, $p = .04$, $\eta^2 = .16$; FRIENDS – $F(1,24) = 5.9$, $p = .02$; $\eta^2 = .19$, FAMILY – $F(1,24) = 6.2$, $p = .02$, $\eta^2 = .21$; or ANXIETY – $F(2,24) = 17.8$, $p < .001$, $\eta^2 = .43$, but there were no group differences or interactions with time. Dating did not improve over time for either group, $F(1,24) = 2.1$, $p = .16$.

3.9. Participant satisfaction

The CSQ-8 has a scale of 1–4, with 4 being excellent, yes definitely, and very satisfied, depending on the item and 1 defined as poor or definitely not. The mean for all items ($M = 3.81$) ranged from a low of 3.33–4.0. The means for each question for parents (P) and the adolescent participant (AP) were: about the quality of service ($P = 3.96$; AP = 3.72), expectations of program met ($P = 3.5$; AP = 3.76), needs met ($P = 3.96$; AP = 3.44), recommend program to others ($P = 3.92$; AP = 3.68), did it help ($P = 3.80$; AP = 3.6), amount of staff help ($P = 3.96$; AP = 3.44), effectiveness ($P = 3.67$; AP = 3.56), overall satisfaction ($P = 3.84$; AP = 3.72), and would you come back if needed ($P = 3.88$; AP = 3.72). There were no group differences for any of the items for either the parent or the participant based on independent *t*-tests. The parent scores generally reflected a slightly higher overall satisfaction than the adolescent participants.

Table 3
Response and remission status at weeks 12, 26 and 52 based on the Children's Depression Rating Scale – Revised (CDRS-R) and Clinical Global Improvement (CGI-I).

CDRS-total score and CGI-I	Exercise	Stretch	Statistic
Week 12			
Responders – CGI-I ≤ 2 and CDRS >50% reduction from baseline	14/14 (100%)	8/12 (67%)	$\chi^2 = 5.52$, df = 1, $p = .019$
Remission – no residual symptoms with CDRS ≤ 28 and CGI-I ≤ 2	12/14 (86%)	6/12 (50%)	$\chi^2 = 3.9$, df = 1, $p = .049$
CDRS-total score and CGI-I	**Exercise**	**Stretch**	**χ^2**
Week 26			
Responders – CGI-I ≤ 2 and CDRS > 50% reduction from baseline	10/10 (100%)	8/10 (80%)	$p = .14$
Remission – no residual symptoms with CDRS ≤ 28 and CGI-I ≤ 2	10/10 (100%)	7/10 (70%)	$p = .06$
Week 52			
Responders – CGI-I ≤ 2 and CDRS > 50% reduction from baseline	7/7 (100%)	7/8 (88%)	$p = .33$
Remission – no residual symptoms with CDRS ≤ 28 and CGI-I ≤ 2	7/7 (100%)	7/8 (88%)	$p = .33$

3.10. Week 26 and 52 follow-up evaluations

From the original 26 subjects who completed the 12 week intervention, 19 (73%) received a six-month follow-up evaluation (10 EXER and 9 STRETCH). The entire EXER group had remitted depressions (100% with no depressive symptoms) by 6 months and the majority had continued to exercise with all reporting that they were basically back to their "normal selves" in terms of school performance, relationships with peers and family and in terms of self-concept. By contrast, for the STRETCH group, 80% had achieved a good clinical response and only 70% had attained remission by six months. The three from the STRETCH group with continued depressive symptoms had been referred for medications and/or psychotherapy and continued in treatment with their personal practitioners.

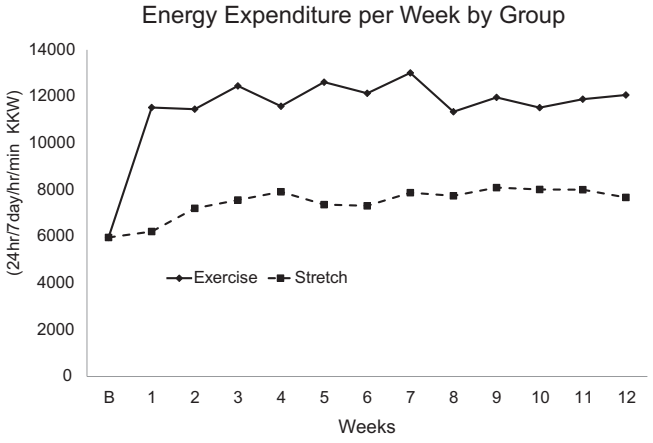


Fig. 3. Energy expenditure in Kilocalories per Kilogram per Week (KKW) by groups.

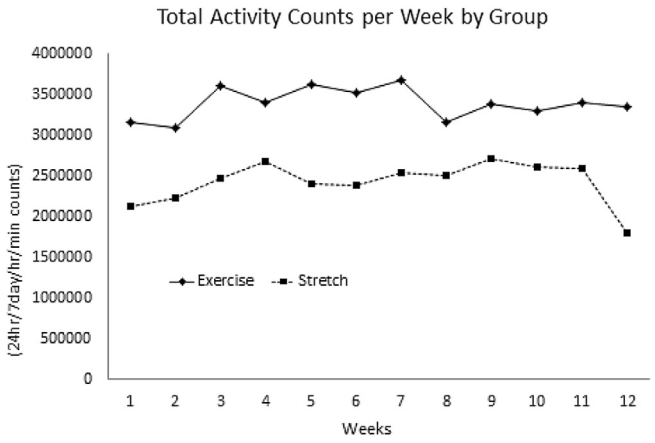
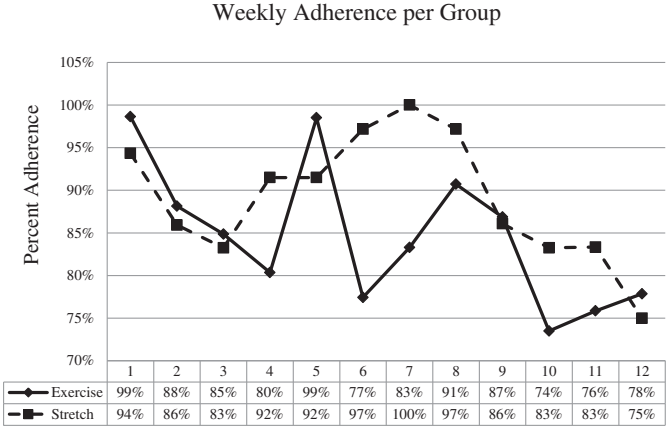


Fig. 4. Total activity counts 24/7 per week by group.

At the end of one year, 15 (58%) returned for a follow-up evaluation (7 EXER and 8 STRETCH). Three had been lost to follow-up since week 26, one from EXER and two from STRETCH. All of those from the EXER group remained remitted (100%) with normal functioning and no return of depression in the interim period. One of the eight from the STRETCH group continued to report mild depressive symptoms. The other 88% had attained remission by the end of 12 months.

4. Discussion

This is the first randomized controlled study of a standardized clinic/home-based aerobic exercise intervention compared to a stretch control condition in mostly non-medicated depressed adolescents meeting DSM-IV-TR MDD criteria. Both the exercise and stretching condition significantly reduced depressive symptoms with the highest clinical response and remission rate to date compared to earlier controlled trials with antidepressants (Brent et al., 2008; Emslie et al., 2008; Emslie, Rush, Weinberg, Kowatch, et al., 1997; March et al., 2004) and adult exercise studies (Blumenthal et al., 2007; Dunn et al., 2005). Exercise did result in a quicker and significantly greater reduction in the CDRS by weeks 6 and 9 resulting in an earlier response than stretching only. Both groups clearly improved by week 12 at which point there were no differences between them on the CDRS measure of depression severity. However, response (100% vs. 67%) and remission (86% vs. 50%) rates were greater in the exercise group compared to the stretching only group. The POMS indicated that in addition to depression improving for both groups, there was also a reduction in anger, fatigue and tension. School performance improved, as did relationships with peers and family based on the SAS-SR. The participants felt better about themselves, and anthropometry measures indicated reduction in triceps, waist, abdominals and hip sizes although weight and BMI did not change. Exercise and stretching routines had an overall positive impact on adolescents who presented with major depressive disorder. Importantly, the intervention was well received with participants reporting high satisfaction with it.



	1	2	3	4	5	6	7	8	9	10	11	12
Exercise	99%	88%	85%	80%	99%	77%	83%	91%	87%	74%	76%	78%
Stretch	94%	86%	83%	92%	92%	97%	100%	97%	86%	83%	83%	75%

Fig. 5. Weekly measure of adherence to exercise prescription by group.

The approach used in this study demonstrated a high adherence rate, indicating that depressed adolescents can be encouraged with structure and reinforcement to engage in various exercise regimes to treat depression. In the adult DOSE study (Dunn et al., 2005), a low dose group did not differ from the stretch group and both groups differed with poorer depression response than the high dose group. DATE, by contrast, found that simple stretching exercises with a low Kcal/kg/week dosage (e.g., <4 KKW) did significantly reduce depression severity progressively over the 12 weeks of intervention from baseline but not as quickly as the more vigorous exercise program (>12 KKW). It had been argued in DOSE that the stretching control group was serving as a placebo condition with only minimal energy expenditure. However, actual energy expenditure and percent maximum heart rates were not monitored or reported for this condition in DOSE, whereas in DATE it was. DATE clearly demonstrated that a stretching condition with good adherence (DOSE had low adherence for stretching), based on the increase in energy expenditure seen overall for the stretching sessions, compared to baseline, may not be a good control or placebo equivalent. The average percent maximum of heart rate as measured during an actual session at the Cooper Institute indicated that the stretching group was achieving 71% maximum versus the 81% for the aerobic exercise group. This is important in indicating that an increase of overall activity that goes along with stretching is not inconsequential, and is an important component of the effectiveness of exercise treatments in general when there is good adherence. The average percent maximum heart rate values were not reported in DOSE (Dunn et al., 2005).

The Dunn et al. (2005) study reports that they had difficulty maintaining adherence in the stretch placebo control group (averaged 42% with a high dropout rate compared to 72% for the aerobic exercise groups) whereas DATE had better adherence for both conditions (STRETCH control 81% and EXER 77% adherence). The subjects of DATE were closely monitored on actual completion of their stretch or exercise sessions and did receive a small monetary compensation for the weekly visit to the CI clinic. This may have improved adherence. The PACES and CSQ-8 self-reports indicated that participants clearly enjoyed the stretching routines comparable to reports for the exercise condition. Another difference that may explain the higher adherence rates was that DATE was designed so that most of the sessions could be done at home or at a site of their preference. By contrast in the adult study (Dunn et al., 2002), all sessions were conducted at the CI, which resulted in additional time and travel commitments that were an added burden. The approach used here argues for the importance of designing exercise intervention studies that address participant burden and makes it easier for them to exercise in a manner that they enjoy with the fewest barriers possible. The participants were allowed flexibility in their choice of exercise for one they enjoyed the most (running, swimming, biking, walking, etc.), as long as the energy expenditure requirements were met. Those assigned to stretching learned new stretching routines each week to minimize boredom. The DATE approach contrasts with some adult studies which limited individuals to stationary bikes or jogging on

Table 4
Baseline anthropometric measures by gender and exercise condition compared to 12 weeks post intervention.

	Condition	Gender	Baseline mean (95% CI)	12 week exit mean (95% CI)
Height (inches)	Exercise	Male	67.4 (64.7–70.1)	67.7 (65.2–70.2)
Gender: $F(1,22) = 4.1$, $p = .05$, $\eta^2 = .16$		Female	62.8 (60.1–65.6)	62.9 (60.4–65.4)
		Total	65.1 (63.2–67.0)	65.3 (63.5–67.0)
	Stretch	Male	65.9 (63.3–68.4)	66.2 (63.9–68.5)
		Female	65.1 (61.5–68.7)	65.4 (62.1–68.7)
		Total	65.6 (63.3–67.7)	65.9 (63.8–67.8)
Weight (pounds)	Exercise	Male	162 (135–190)	167 (140–195)[a]
Weeks × Gender: $F(1,21) = 5.4$, $p = .026$, $\eta^2 = .21$		Female	161 (131–190)	160 (160–190)
		Total	162 (141–182)	164 (143–184)
	Stretch	Male	131 (105–157)	135 (109–161)[a]
		Female	157 (121–194)	156 (120–193)
		Total	140 (122–166)	142 (123–168)
Waist (inches)	Exercise	Male	31.4 (27.6–35.2)	32.0 (27.9–36.1)
Weeks × Gender: $F(1,21) = 6.1$, $p = .02$, $\eta^2 = .22$		Female	35.0 (31.0–39.1)	34.7 (30.3–39.1)
		Total	33.2 (30.4–36.0)	33.4 (30.4–36.4)
	Stretch	Male	29.3 (25.8–32.8)	29.7 (25.9–33.5)
		Female	35.1 (30.7–40.1)	31.8 (26.4–37.2)[a]
		Total	32.2 (29.1–35.2)	30.8 (27.5–34.1)
Hip (inches)	Exercise	Male	37.0 (33.2–40.7)	37.0 (33.2–40.9)
Gender: $F(1,21) = 4.4$, $p = .05$, $\eta^2 = .17$		Female	41.7 (37.6–45.8)	40.1 (35.9–44.2)[a]
Weeks × Group: $F(1,21) = 7.3$, $p = .01$, $\eta^2 = .26$		Total	39.2 (36.6–42.1)	38.4 (35.7–41.4)
Weeks × Group × Gender: $F(1,21) = 6.6$, $p = .02$, $\eta^2 = .24$	Stretch	Male	33.8 (30.3–37.4)	34.0 (330.4–37.6)
		Female	36.9 (31.9–41.9)	39.6 (34.5–44.7)[a]
		Total	34.8 (32.3–38.4)	35.8 (33.7–39.9)
Triceps skin fold (mm)	Exercise	Male	18.9 (12.8–24.9)	16.7 (10.7–22.6)
Gender: $F(1,8) = 16$, $p = .001$, $\eta^2 = .47$		Female	28.9 (23.4–34.4)	27.0 (21.6–32.4)
Weeks × Group: $F(1,18) = 6.9$, $p = .02$, $\eta^2 = .28$		Total	23.9 (19.8–28.0)	21.8 (17.8–25.8)
	Stretch	Male	12.4 (7.6–17.1)	13.2 (8.5–17.9)
		Female	22.0 (14.2–29.8)	27.0 (19.3–34.7)
		Total	17.2 (12.6–21.8)	20.1 (15.6–24.6)
Abdominals skin fold (mm)	Exercise	Male	24.4 (17.1–31.7)	22.1 (13.6–30.6)
Gender: $F(1,18) = 5.8$, $p = .03$, $\eta^2 = .24$		Female	32.1 (25.4–38.8)	29.5 (21.7–37.3)
Weeks × Group: $F(1,18) = 4.8$, $p = .04$, $\eta^2 = .21$		Total	28.2 (23.3–33.2)	25.8 (20.0–31.6)
	Stretch	Male	15.8 (10.0–21.6)	16.6 (9.8–23.3)
		Female	24.7 (15.2–34.1)	28.2 (17.2–39.2)
		Total	20.2 (14.7–25.8)	22.4 (15.9–28.8)
Thighs skin fold (mm)	Exercise	Male	26.7 (16.6–36.9)	20.4 (11.4–29.4)
Gender: $F(1,18) = 12.1$, $p = .003$, $\eta^2 = .40$		Female	39.4 (30.2–48.6)	37.5 (29.2–45.8)
Weeks × Group: $F(1,18) = 9.1$, $p = .007$, $\eta^2 = .34$		Total	33.1 (26.2–39.9)	29.0 (22.8–35.1)
Weeks × Gender: $F(1,18) = 4.3$, $p = .05$, $\eta^2 = .19$	Stretch	Male	18.9 (10.8–26.9)	19.9 (12.7–27.0)
		Female	30.3 (17.3–43.4)	39.5 (27.8–51.2)
		Total	24.6 (16.9–32.3)	29.7 (22.8–36.5)
BMI	Exercise	Male	24.8 (19.8–29.8)	25.4 (20.2–30.5)
Gender : $F(1,22) = 4.95$, $p = .037$, $\eta^2 = .18$		Female	32.4 (27.4–37.4)	32.4 (27.3–37.5)
		Total	28.6 (25.1–32.2)	28.9 (25.3–32.5)
	Stretch	Male	21.2 (16.5–25.9)	21.7 (16.9–26.5)
		Female	25.9 (19.3–32.6)	25.6 (18.8–32.4)
		Total	22.8 (19.5–27.6)	23.0 (19.5–27.8)
Percent body fat	Exercise	Male	16.4 (11.5–21.2)	15.4 (10.4–20.5)
Gender: $F(1,21) = 43.0$, $p = .000$, $\eta^2 = .67$		Female	32.4 (27.1–37.7)	32.0 (26.5–37.4)
		Total	23.8 (20.8–28.0)	23.1 (20.0–27.4)
	Stretch	Male	10.7 (6.1–15.2)	11.1 (6.4–15.9)
		Female	27.0 (20.6–33.4)	28.8 (22.2–35.6)
		Total	16.1 (14.9–22.8)	17.0 (15.9–24.1)
Bone density (×10)	Exercise	Male	10.6 (10.5–10.7)	10.6 (10.5–1.08)
Gender: $F(1,21) = 42.0$, $p = .000$, $\eta^2 = .67$		Female	10.2 (10.1–11.4)	10.2 (10.2–10.4)
		Total	10.5 (10.4–10.5)	10.5 (10.4–1.05)
	Stretch	Male	10.7 (10.6–10.9)	10.7 (10.6–10.8)
		Female	10.4 (10.2–10.5)	10.3 (10.2–10.5)
		Total	10.6 (1.05–1.07)	10.6 (1.05–10.6)

[a] Indicates which gender had a significant change from baseline measure.

treadmills and required that every session be completed at the study site. Other adult studies have utilized exercise interventions which encouraged people to exercise after providing recommendations but did not objectively measure actual physical activity or have interventions in place to address adherence issues. Unlike DATE where blinded clinician ratings and 24/7 activity monitoring were included, some studies have relied on less reliable self-reports for depression and adherence.

4.1. Response, remission rates and effect sizes

To put the clinical response and remission rate results of DATE in perspective, in the first acute phase fluoxetine study of depressed youth (Emslie, Rush, Weinberg, Kowatchm, et al., 1997), 56% of those on fluoxetine versus 33% of those on placebo rated themselves as "much" or "very much" improved at study exit. These results were replicated in a large, multi-site trial with 52.3% for fluoxetine vs. 36.8% for placebo (Emslie et al., 2002) with similar findings for another multisite study (March et al., 2004). Meta-analyses of adult studies ($n = 27$) report an overall 46.4% response rate to fluoxetine and a 24.7% response rate to placebo (Panel, 1993a, 1993b). Clearly, fluoxetine is an effective treatment but with a lot of non- or partial responders, residual symptoms and side effects (AACAP, 2007; Fava et al., 2002; Hughes et al., 2007; Kennard et al., 2006). By contrast, using the same criteria, the EXER condition in this study had a 100% clinical response rate and STRETCH routines had a 67% clinical response.

Total remission was 86% for EXER and 50% for STRETCH in DATE. There were no medication side effects and/or other side effects (adverse events) reported for EXER or STRETCH. A more recent fluoxetine study found that relapse occurred more frequently in participants randomized to placebo (36/52; 69.2%) than fluoxetine (21/50; 42.0%) ($p = .009$) following a 12-week open trial (Brent et al., 2008; Emslie et al., 2008). DATE had only two fail to remit by six and 12 months following treatment. This is an important finding since adolescents with more symptoms remaining at the end of 12 weeks of treatment are less likely to be remitted at 18 and 36 weeks post initiation of treatment and are more likely to relapse. In addition, faster rate of response in the acute phase of treatment is associated with more complete improvement, with those responding to treatment within the first four weeks being more likely to achieve remission after 12 weeks of treatment (Brent et al., 2008; Tao et al., 2009). The DATE EXER group had a rapid rate of response compared to STRETCH and comparable to medication trials which argues for a safe non-medication intervention. The six and twelve month follow-up visits found no relapses in DATE for any of those that had good clinical response and remission at the follow-up visits. Many indicated that they had continued with their various exercise and/or stretching regimes.

4.2. Strengths and limitations

4.2.1. Strengths

The current study was designed to address previous criticisms and limitations of adolescent exercise studies, most importantly; no previous study had used a blinded randomized control trial (Dunn et al., 2005). Second, DATE improved upon prior studies by requirement of strict DSM-IV-TR diagnostic criteria for major depressive disorder ascertained by structured diagnostic interviewing (KSADS-PL). Third, DATE used a standardized entry and exclusion criteria similar to previous randomized controlled antidepressant pharmaceutical trials and depression outcome measures using the same gold-standard instruments (Brent et al., 2008; Emslie et al., 2008; Hughes et al., 2007; March et al., 2004). Fourth, a blinded clinician rater derived the CDRS to allow comparison to established depressed adolescent research data. Additional standardized measures included weekly self-report by both the participant and their parents. Fifth, instead of relying on self-reports or exercise diaries for reports of exercise sessions completed, the participants were required to wear Actical KKW monitoring devices 24 h/day for the 84 days of the study (not including prior to baseline) that captured the reported exercise session as well as activity for the rest of the day. This is the first reported study that included 24 h/day monitoring of energy expenditure beyond just that for the prescribed exercise session. The weekly monitoring allowed for adherence checks both for wearing the Actical device and sessions completed which allowed for problem-solving with the participant and the parents. Provision of a small monetary reinforcement for maintaining the prescribed exercise sessions may also have helped in achieving high adherence rates and minimizing dropouts compared to earlier studies.

Prior to this study there were no exercise studies of depressed adolescents (who met DSM-IV-TR MDD criteria) that had established effect sizes for the various interventions and measures typically used to study depression. This study included many different measures and positive findings with large effect sizes with a relatively small sample of subjects (c.f. with adult studies with similar small ns). These preliminary effect sizes for various measures will help to inform future studies for determining power and sample sizes (Ferguson, 2009). On the other hand, there may have been some missed findings due to Type II error. In addition, few exercise studies exist with longitudinal follow-up data. Continued wellness in the participants at follow-up six and twelve months later speaks to the importance of including exercise as a viable treatment alternative to medication, which typically has a fairly high relapse rate (Emslie, Rush, Weinberg, Kowatchm, et al., 1997; Kennard et al., 2006; Tao et al., 2009).

4.2.2. Limitations

The DATE study had a relatively small sample size, even though many of the observed effect sizes (partial eta-squared) were large. The cell sizes of the current study are actually comparable to that of large studies (Dunn et al., 2005). The study was limited to adolescents and no cost measures were included. DATE findings for adolescents provide support for a future study of younger children to increase generalization. It was not possible to blind the participants or the parents to the exercise condition although the clinician raters were. And given that these were clinically depressed adolescents with school, family and social impairment, we were unable to ethically have a true no-treatment (or waitlist) control group to establish an equivalent "spontaneous recovery" rate or treatment as usual as done in adult depression studies. If one were to assume a 30%–40% placebo response rate based on prior adolescent antidepressant trials, the response rate for exercise intervention remains a remarkably strong finding and argues that the DATE study findings were not merely a placebo effect. A future study should compare exercise to pharmaceutical, cognitive behavior therapy, combinations of treatment, and its role in augmentation studies for adolescents.

5. Conclusions

Depressed adolescents will participate in a structured exercise intervention without medication, and such treatment appears to offer a viable alternative for those seeking a non-medication intervention that appears to be beneficial. DATE findings support that six-twelve weeks of either aerobic exercise (12 KKW) or a series of weekly stretching routines (less than 4 KKW) can be effective in treating depressed adolescents with the magnitude of the effect being stronger for the former when there are procedures

in place and good support for achieving adherence. There are large observed effect sizes for the clinical outcome measures when adherence occurs in contrast to those studies where adherence is poor or not monitored. The results appear to match, if not exceed, prior research with antidepressants while avoiding the side effects, costs or risks associated with antidepressant medication. Future research with youth should compare exercise to an established antidepressant in a randomized control trial similar to an earlier study with adults (Blumenthal et al., 2007). Exercise interventions have promise as a non-medication intervention for depressed adolescents and further randomized control trials are merited.

Financial disclosure

Carroll Hughes and Graham Emslie are consultants for BioBehavioral Diagnostics, Inc. Graham Emslie has received research support from BioBehavioral Diagnostics Inc., BioMarin, Eli Lilly, Forest, GlaxoSmithKline, and Somerset; has served as a consultant for Bristol-Myers Squibb, Eli Lilly, Forest, GlaxoSmithKline, INC Research Inc., Lundbeck, Pfizer, Seaside Therapeutics Inc., Shire, Valeant, and Wyeth; and has been on the Speakers Bureau for Forest. Reprint requests to: Carroll W. Hughes, Ph.D., ABPP, Department of Psychiatry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, Texas, 75390-8589.

Funding support

This study was supported by a grant entitled Exercise for Depressed youth (R34 MH075762) from the National Institutes of Health (Principal Investigator: Carroll W. Hughes).

The trial was registered 2/18/2009 with ClinicalTrials.gov with the following website address: http://clinicaltrials.gov/ct2/show/NCT00847457?term=ADOLESCENT+DEPRESSION+DATE&rank=1 and is now reported as completed.

Acknowledgments

The support, guidance and encouragement early on from Drs. Barbara Burns, Russell Pate, and Timothy Church as external consultants were invaluable. A number of UT Southwestern Staff also provided guidance and assistance early on including Drs. Madhukar Trivedi, Tracey Greer, Betsy Kennard, A. John Rush, and project administrative support from Taryn Mayes and Jeanne Rintelmann. The project would not have been possible without all the contributions of support staff at the Cooper Institute in Dallas, Texas, associated with the project over the duration of the study: Melba Morrow, Tyson Bain, Judy Dubreuil, Dr. Joseph Cleaver, Lucille Marcoux, Susan Devers, Erica Howard, Tiffany Gearhart, Patrick Greak, Christina Duncan, Andrew Vidales. And most importantly, we would like to thank the adolescents and families that participated in this study.

References

AACAP. (2007). Practice parameter for the assessment and treatment of children and adolescents with Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(7), 894–921.

Aydemir, O., Yalcin, E. S., Aksaray, S., Kisa, C., Yildirim, S. G., & Uzaby, T. (2006). Brain-derived neurotrophic factor (BDNF) changes in the serum of depressed women. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 30*, 1256–1260.

Babyak, M., Blumenthal, J. A., Herman, S., Khatri, P., Doriswamy, M., Moore, K., et al. (2000). Exercise treatment for major depression: maintenance of therapeutic benefit at 10 months. *Psychosomatic Medicine, 62*, 633–638.

Bernstein, I. H., Rush, A. J., Trivedi, M. H., Hughes, C. W., Macleod, L., Witte, B. P., et al. (2010). Psychometric properties of the Quick Inventory of Depressive Symptomatology in adolescents. *International Journal of Methods in Psychiatric Research, 19*, 185–194.

Biddle, S. J. H. (2001). Adherence to sport and physical activity in children and youth. In S. Bull (Ed.), *Adherence issues in sport and exercise*. New York.

Birmaher, B., & Brent, D. A. (1998). Work Group on Quality Issues. Practice parameters for the assessment and treatment of children and adolescents with depressive disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 37*, 63S–83S.

Blumenthal, J. A., Babyak, M. A., Doraiswamy, P. M., Watkins, L., Hoffman, B. M., Barbour, K. A., et al. (2007). Exercise and pharmacotherapy in the treatment of major depressive disorder. *Psychosomatic Medicine, 69*, 587–596. http://dx.doi.org/10.1097/PSY.0b013e318148c19a.

Brent, D. A., Emslie, G. J., Clarke, G., Wagner, K. D., Asarnow, J. R., Keller, M., et al. (2008). Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression. *Journal of the American Medical Association, 299*, 901–913.

Brent, D. A., Kolko, D. J., Birmaher, B., Baugher, M., & Bridge, J. (1999). A clinical trial for adolescent depression: predictors of additional treatment in the acute and follow-up phases of the trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*, 263–270.

Curran, S. L., Andrykowski, M. A., & Studts, J. L. (1995). Short Form of the Profile of Mood States (POMS-SF). *Psychological Assessment, 7*, 80–83.

Curry, J., Silva, S., Rohde, P., Ginsburg, G., Kratochivil, C., Kirchner, S. A., et al. (2011). Recovery and recurrence following treatment for adolescent major depression. *Archives of General Psychiatry, 68*, 263–269.

Dunn, A. L., & Dishman, R. K. (1991). Exercise and the neurobiology of depression. *Exercise Sport Science Review, 19*, 41–48.

Dunn, A. L., Marcus, B. H., Kampert, J. B., Garcia, M. E., Kohl, H., & Blair, S. N. (1999). Comparison of lifestyle and structured interventions to increase physical activity and cardiorespiratory fitness: a randomized trial. *Journal of American Medical Association, 28*(14), 327–334.

Dunn, A. L., Trivedi, M. H., Kampert, J. B., Clark, C. G., & Chambliss, H. O. (2002). The DOSE study: a clinical trial to examine efficacy and dose response of exercise as treatment for depression. *Controlled Clinical Trials, 23*, 584–603.

Dunn, A. L., Trivedi, M. H., Kampert, J. B., Clark, C. G., & Chambliss, H. O. (2005). Exercise treatment for depression: efficacy and dose response. *American Journal of Preventive Medicine, 28*, 1–8.

Dunn, A. L., Trivedi, M. H., & O'Neal, H. A. (2001). Physical activity dose-response effects on outcomes of depression and anxiety. *Medicine and Science in Sports and Exercise, 33*(Suppl. 6), S587–S597.

Dunn, A. L., & Weintraub, P. (2008). Exercise in the prevention and treatment of adolescent depression: a promising but little researched intervention. *American Journal of Lifestyle and Medicine, 2*, 207–518.

Dwivedi, Y., Mondal, A. C., Rizavi, H. S., & Conley, R. R. (2005). Suicide brain is associated with decreased expression of neurotrophins. *Biological Psychiatry, 58*, 315–324.

Emslie, G. J., Heiligenstein, J. H., Wagner, K. D., Hoog, S. L., Ernest, D. E., Brown, E., et al. (2002). Fluoxetine for acute treatment of depression in children and adolescents: a placebo-controlled, randomized clinical trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*, 1205–1215.

Emslie, G. J., Kennard, B. D., Mayes, T. L., Nightingale-Teresi, J., Carmody, T., Hughes, C. W., et al. (2008). Fluoxetine versus placebo in preventing relapse of major depression in children and adolescents. *American Journal of Psychiatry, 165*(4), 459–467.

Emslie, G. J., Rush, A. J., Weinberg, W. A., Gullion, C. M., Rintelmann, J., & Hughes, C. W. (1997). Recurrence of major depressive disorder in hospitalized children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 785–792.

Emslie, G. J., Rush, A. J., Weinberg, W. A., Kowatch, R. A., Hughes, C. W., Carmody, T., et al. (1997). A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Archives of General Psychiatry, 54*, 1031–1037.

Fava, G. A., Fabbri, S., & Sonino, N. (2002). Residual symptoms in depression: an emerging therapeutic target. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 26*, 1019–1027.

Ferguson, C. J. (2009). An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*, 532–538.

Ferris, L. T., Williams, J. S., & Shen, C. L. (2007). The effect of acute exercise on serum brain-derived neurotrophic factor levels and cognitive function. *Medical Science and Sport Exercise, 39*, 728–734.

Gonul, A. S., Akdeniz, F., Tanelie, F., Donat, O., Eker, C., & Vahip, S. (2005). Effect of treatment on serum brain-derived neurotrophic factor levels in depressed patients. *European Archives of Psychiatry and Clinical Neuroscience, 255*, 381–386.

Green, S. B., & Salkind, N. J. (2011). *Using SPSS for Windows and Macintosh: Analyzing and understandting data* (6th ed.). New York: Prentice Hall.

Guy, W. (1976). *Clinical global improvement scale*. Rockville, MD: US Department of Health, Education, and Welfare.

Huang, T. L., Lee, C. T., & Liu, Y. L. (2007). Serum brain-derived neurotrophic factor levels in patients with major depression. *Journal of Psychiatric Research, 42*, 521–525.

Hughes, C. W., Emslie, G. J., Crismon, M. L., Posner, K., Birmaher, B., Ryan, N., et al. (2007). Texas Children's medication Algorithm project: update from Texas Consensus Conference Panel on medication treatment of childhood major depressive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(6), 667–686.

Hughes, C. W., Trivedi, M. H., Cleaver, J., Greer, T. L., Emslie, G. J., Kennard, B., et al. (2009). DATE: depressed adolescents treated with exercise: study rationale and design for a pilot study. *Mental Health and Physical Activity, 2*(2), 76–85.

Judd, L. L., Paulua, M. J., Schettler, P. J., Akiskal, H. S., Endicott, J., Leon, A. C., et al. (2000). Does incomplete recovery from first lifetime major depressive episode herald a chronic course of illness? *American Journal of Psychiatry, 157*, 1501–1504.

Karege, F., Perret, G., Bondolifi, G., Schwald, M., Bertschy, G., & Aubry, J. M. (2002). Decreased serum brain-derived neurotrophic factor levels in major depressed patients. *Psychiatry Research, 109*, 143–148.

Kaufman, J., Birmaher, B., Brent, D. A., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 980–988.

Keller, S., Sarchiapone, M., Zarrilli, F., Videtic, A., Ferraro, A., Carli, V., et al. (2010). Increased BDNF promoter methylation in the Wernicke Area of suicide subjects. *Archives of General Psychiatry, 67*(33), 258–267.

Kendzierski, D., & DeCarlo, K. J. (1991). Physical activity enjoyment scale: two validation studies. *Journal of Sport & Exercise Psychology, 13*, 50–64.

Kennard, B. D., Silva, S., Vitiello, B., Curry, J., Kratochivil, C., Simons, A., et al. (2006). Remission and residual symptoms after acute treatment of adolescents with major depressive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*, 1404–1411.

Kratochvil, C. J., Vitiello, B., & Walkup, J. (2006). SSRIs in pediatric depression: is the balance between benefits and risk favorable? *Journal of Child and Adolescent Psychopharmacology., 16*, 11–14.

Larun, L., Nordheim, L. V., Ekeland, E., Hagan, K. B., & Heian, F. (2006). Exercise in prevention and treatment of anxiety and depression among children and young people. *Cochrane Database of Systemic Reviews, 3*, 1–40.

Lawlor, D. A., & Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomized controlled trials. *British Medical Journal, 322*, 763–767.

Lin, E. H., Katon, W. J., VonKorff, M., Russo, J. E., Simon, G. E., Bush, T. M., et al. (1998). Relapse of depression in primary care. Rate and clinical predictors. *Archives of Family Medicine, 7*, 443–449.

March, J. S., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J., et al. (2004). Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adoelscents with Depression Study (TADS) randomized control trial. *Journal of the American Medical Asssociation, 292*(7), 807–820.

Mead, G. E., Morley, W., Campbell, P., Greig, C. A., McMurdo, M., & Lawlor, D. A. (2009). Exercise for depression (review). *Cochrane Database of Systematic Reviews, 3*, 1–61. http://dx.doi.org/10.1002/14651858.CD004366.pub4.

Meeusen, R., Piacentini, M. F., & DeMeirleir, K. (2001). Brain microdialysis in exercise research. *Sports Medicine, 31*, 965–983.

Morss, G. M., Jordan, A. N., Skinner, J. S., Dunn, A. L., Church, T. S., Earnest, C. P., et al. (2004). Dose response to exercise in women aged 45–75 yr (DREW): design and rationale. *Medicine and Science in Sports and Exercise, 36*(2), 336–344.

Nabkasorn, C., Miyai, N., Sootmongkol, A., Junprasert, S., Yamamoto, H., Arita, M., et al. (2006). Effects of physical exercise on depression, neuroendocrine stress hormones and physiological fitness in adolescent females with depressive symptoms. *European Journal of Public Health, 16*(2), 179–184. http://dx.doi.org/10.1093/eurpub/cki159.

Nemeroff, C. B., Kalali, A., & Keller, M. B. (2007). Impact of publicity concerning pediatric suicidality data on physician practice patterns in the United States. *Archives of General Psychiatry, 64*(4), 466–472.

Nguyen, T., Attkisson, C., & Stegner, B. (1983). Assessment of patient satisfaction development and refinement of a service evaluation. *Evaluation and Program Planning, 6*(3), 299–313.

O'Neal, H. A., Dunn, A. L., & Martinsen, E. W. (2000). Depression and exercise. *International Journal Sports Psychology, 31*, 110–135.

Oeland, A. M., Laessoe, U., Olesen, A. V., & Munk-Jergensen, P. (2010). Impact of exercise on patients with depression and anxiety. *Nordic Journal of Psychiatry, 4*(64), 210–217.

Olfson, M., Marcus, S. C., & Shaffer, D. (2006). Antidepressant drug therapy and suicide in severely depressed children and adults. *Archives of General Psychiatry, 63*, 865–872.

Paluska, S. A., & Schwenk, T. L. (2000). Physical activity and mental health: current concepts. *Sports Medicine, 29*(3), 167–180.

Panel, Depression Guideline. (1993a). Clinical practice guideline. In *Detection and diagnosis: Vol. 1. Depression in primary care*. U.S. Department of Health and Human Services-Agency for Health Care Policy Research.

Panel, Depression Guideline. (1993b). In *Detection and diagnosis: Vol. 2. Depression in primary care*. Treatment. Paper presented at the Clinical Practice Guidelines, Rockville, MD.

Penedo, F. J., & Dahn, J. R. (2005). Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current Opinions in Psychiatry, 18*(2), 189–193.

Poznanski, E. O., & Mokros, H. B. (1996). *Children's Depression Rating Scale, Revised (CDRS-R) manual*. Los Angeles: Western Psychological Services Publishers and Distributors.

Rimer, J., Dwan, K., Lawlor, D. A., Greig, C. A., McMurdo, M., & Mead, G. E. (2012). Exercise for depression. *Cochrane Database of Systematic Reviews, 10*. (Art. No.: CD04366).

Rothon, C., Edwards, P., Bhul, K., Viner, R. M., Taylor, S., & Stansfeld, S. A. (2010). Physical activity and depressive symptoms in adolescents: a prospective study. *BMC Medicine, 28*(6), 32.

Rush, A. J., Bernstein, I. H., & Trivedi, M. H. (2006). An evaluation of the Quick Inventory of Depressive Symptomatology and the Hamilton Rating Scale for Depression: a STAR*D report. *Biological Psychiatry, 59*(6), 493–501.

Rush, A. J., Carmody, T. J., & Ibrahim, H. M. (2006). Comparison of self-report and clinician ratings on two inventories of depressive symptomatology. *Psychiatric Services, 57*(6), 829–837.

Russo-Neustadt, A., Beard, R. C., & Cotman, C. W. (1999). Exercise, antidepressant medications, and enhanced brain derived neurotrophic factor expression. *Neuropsychopharmacology, 21*, 679–682.

Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P. J., Fisher, P., & Bird, H. (1983). A children's global assessment scale (C-GAS). *Archives of General Psychiatry, 40*, 1228–1231.

Siuciak, J. A., Lewis, D. R., Wiegand, S. J., & Lindsay, R. M. (1997). Antidepressant-like effect of brain-derived neurotrophic factor (BDNF). *Pharmacology Biochemistry & Behavior, 56*, 131–137.

Sund, A. M., Larsson, B., & Wichstrom, L. (2011). Role of physical and sedentary activities in the development of depressive symptoms in early adolescence. *Social Psychiatry and Psychiatric Epidemiology, 46*, 431–441.

Szabo, A., Billett, E., & Turner, J. (2001). Phenylethylamine, a possible link to the antidepressant effects of exercise? *British Journal of Sports Medicine, 35*, 342–343.

Taliaferro, L. A., Rienzo, B. A., Pigg, R. M., Miller, M. D., & Dodd, V. J. (2009). Associations between physical activity and reduced rates of hopelessness, depression, and suicidal behavior among college students. *Journal of the American College of Health, 57*(4), 427–436.

Tao, R., Emslie, G. J., Mayes, T. L., Nakonezny, P. A., Kennard, B. D., & Hughes, C. W. (2009). Early prediction of acute depressant treatment response and remission in pediatric major depressive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(1), 71–78.

Trivedi, M. H., Greer, T. L., Grannemann, B. D., Church, T. S., Galper, D. I., Sunderajan, P., et al. (2006). TREAD: Treatment with Exercise Augmentation for Depression: study rationale and design. *Clinical Trials, 3*(3), 291–305.

Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry, 33*(9), 1111–1115.

Wilkinson, L. (1999). Statistical methods in psychology journals: guidelines and explanations. *American Psychologist, 54*, 594–604.

Zarrilli, F., Angiolillo, A., Castaldo, G., Chairiotti, L., Keller, S., Sacchetti, S., et al. (2009). Brain derived neurotrophic factor (BDNF) genetic polymorphism (Val66Met) in suicide: a study of 512 case. *American Journal of Medical Genetics and Neuropsychiatry of Genetics, 1508*(4), 599–600.