Physical exercise for late-life depression: Effects on symptom dimensions and time course

Physical exercise for late-life depression: effects on symptom dimensions and time course

Martino Belvederi Murri [1], Pantaleimon Ekkekakis [2], Marco Menchetti [3], Francesca Neviani [4], Fausto Trevisani [5], Stefano Tedeschi [6], Pasqualino Maietta Latessa [7], Erika Nerozzi [8], Giuliano Ermini [9] Donato Zocchi [9] Salvatore Squatrito [10], Giulio Toni [11], Aderville Cabassi [6], Mirco Neri [4], Stamatula Zanetidou [12], Mario Amore [1]

1. Section of Psychiatry, Department of Neuroscience, Oftalmology, Genetics and Infant-Maternal Science, University of Genoa, Genoa, Italy

2. Department of Kinesiology, Iowa State University, USA

3. Department of Medical and Surgical Sciences, University of Bologna, Italy

4. Department of Geriatrics, Nuovo Ospedale Civile S. Agostino Estense, Modena and Reggio Emilia University, Modena, Italy

5. Local Health Agency (AUSL), Bologna, Italy

6. Department of Clinical and Experimental Medicine, Endocrinology of Aging Unit, University of Parma, Italy

7. Department for Life Quality Studies, University of Bologna, Italy

8. Department of Pharmacy, Biotechnology and Motor Sciences, University of Bologna, Italy

9. Primary Care Physician, Bologna, Italy

10. Department of Biomedical and Neuromotor Sciences, Section of Human and General Physiology, University of Bologna, Italy.

11. Cardiology Unit, Hospital S. Sebastiano, Correggio, Italy

12. Consultation Liaison Psychiatry Service, Department of Mental Health, Bologna, Italy

* *Corresponding author:*

Martino Belvederi Murri, Section of Psychiatry, *Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health*, University of Genoa.

Largo Rosanna Benzi, 10, Genova, 16132, Italy;

Email: martino.belvederi@gmail.com; Telephone: +39 333 5248720

Abstract

Background. Physical exercise is increasingly recognized as a treatment for major depression, even among older patients. However, it is still unknown which depressive symptoms exercise affects most, (e.g. somatic vs. affective) and the timing of its effects. Thus, the aim of this study was to examine the changes of depressive symptoms after treatment with exercise.

Methods. We analyzed data from the SEEDS study, a trial comparing the antidepressant effectiveness of sertraline (S) and sertraline plus exercise (S+EX). Exercise was delivered thrice weekly in small groups and monitored by heart rate meters. Patients with late life depression (n=121) were assessed at baseline, 4, 8, 12 and 24 weeks with the Hamilton Depression Scale. Scores of affective, vegetative, anxiety and agitation/insight factors were analyzed using Multilevel Growth Curve Models and sensitivity analyses (multiple imputation).

Results. Compared with the S group, patients in the S+EX group displayed significantly greater improvements of the affective symptom dimension (total effect size=0.79) with largest changes in the first 4 weeks and last 12 weeks. Improvements were mainly driven by depressed mood and psychomotor retardation.

Limitations. Sample size; lack of an exercise only treatment arm

Conclusions. Adding exercise to antidepressant drug treatment may offer significant advantages over affective symptoms of depression, rather than somatic symptoms or other dimensions of depression. Compared with standard antidepressant treatment, clinical advantages should be expected both at an early (first 4 weeks) and later stage (after 12 weeks).

Keywords: depression; antidepressants; exercise; mood; psychomotor retardation; affective

Word count: 3031

Running title: Exercise for old age depression: effects on symptoms

1. Introduction

Physical exercise is increasingly recognized as an effective treatment for late-life depression (Heinzel et al., 2015; Schuch et al., 2016b) but its specific effects on symptom dimensions are still largely unknown.

Major depression is regarded as one of the most prevalent and debilitating healthcare problems worldwide, with dire consequences for individuals, families, and society as a whole (Alexopoulos, 2005). Its clinical presentation is highly variable, including both "core" depressive symptoms, such as low mood and reduced interest for activities, and somatic symptoms, such as sleep and appetite changes. In late life, the diagnosis and treatment of depression are further complicated by specific neurobiological features and by the cooccurrence of physical illnesses (Naismith et al., 2012). These factors have a profound impact on response to treatments, which is lower than among younger individuals (Alexopoulos, 2005), and on its clinical presentation, which more frequently includes somatic symptoms, apathy, and psychomotor retardation (Groeneweg-Koolhoven et al., 2017; Haigh et al., 2017; Hegeman et al., 2015; Naismith et al., 2012). Notably, differences in the clinical presentation may correspond to distinct pathophysiological mechanisms and differential responses to treatment (Drysdale *et al.*, 2017;Li *et al.*, 2016). Likewise, different treatments may have differential impact on symptom dimensions (Uher et al., 2012). Since late-life depression is characterized by suboptimal drug response and poor outcomes in the real-world clinical practice, there is an urgent need to improve the present understanding of novel therapeutic strategies and the mechanisms via which they influence the clinical features of depression (Alexopoulos, 2005).

Physical exercise is increasingly recognized as an effective tool for the management of depression. Exercise participation has been found to substantially reduce the severity of depressive symptoms (Ekkekakis, 2015; Schuch et al., 2016b), is well tolerated (Stubbs et al., 2016b), increases physical fitness (Stubbs et al., 2016a), and is among patient preferred treatment options (Luck-Sikorski et al., 2017). Studies on late-life depression broadly confirm the positive results observed among younger samples (Heinzel et al., 2015; Schuch et al., 2016b). However, the question remains whether, among older patients, the efficacy of exercise is due to "non-specific" effects on somatic symptoms (i.e. improvements of sleep, appetite, tiredness) or it also encompasses improvements in "core" depressive symptoms, such as depressed mood and lack of interest (Ekkekakis and Belvederi Murri, 2017). Considerable evidence indicates that physical exercise has the potential to both improve mood in the short term (Ekkekakis et al., 2011) and stimulate long-terms antidepressant mechanisms, such as neurogenesis (Ekkekakis, 2013; Kerling et al., 2017; Schuch et al., 2016a). Therefore, the evaluation of the effects of exercise across different symptom-dimensions over multiple time points may assist

clinicians both in terms of monitoring the response to treatment and in guiding the prognosis (Iniesta et al., 2016; Uher et al., 2009).

The aim of the present study was to evaluate the effects of a program of physical exercise on symptom dimensions of late-life depression, taking in account the timing of these changes. The study was based on data from the SEEDS (Safety and Efficacy of Exercise for Depression in Seniors) trial, which randomized patients with late-life depression to antidepressant drugs or antidepressants plus structured physical exercise (Belvederi Murri *et al.*, 2015). Our hypothesis was that patients receiving exercise in addition to antidepressants would display greater and earlier improvements in "core" depressive symptoms, compared to patients receiving only antidepressants.

2. Methods

2.1 The SEEDS study

SEEDS was a randomized trial examining the effectiveness of two exercise interventions, combined with standard antidepressant treatment, against antidepressant treatment alone. Details on the study protocol are available in a previous report (Belvederi Murri et al., 2015). Briefly, the study enrolled 121 participants aged 65 – 85 years diagnosed with Major Depression (DSM-IV TR criteria) from four centers in the region of Emilia Romagna, Italy. Participants were selected by Primary Care Physicians (PCPs) and interviewed by psychiatrists in the context of a liaison program between the Mental Health and Primary Care Departments. Other selection criteria included: a score of 18 or higher on the 17-item Hamilton Depression Rating Scale (HAM-D), being sedentary (not meeting the recommended levels of physical activity for older adults (Nelson et al., 2007)), absence of other axis I diagnoses, substance or alcohol abuse, severe or unstable physical illness that would prevent them from exercising (e.g. severe cardiovascular disease, osteoarthritis, uncontrolled diabetes, major neurological disorders, severe respiratory disease) and cognitive impairment (Mini Mental State Examination score of 24 or higher). Participants were given information on the study interventions and on the effects of exercise during meetings with their PCPs and study staff.

Of 177 participants who were referred by PCPs for evaluation, 121 fulfilled inclusion criteria and were randomized to study interventions. Participants were assigned to 1) sertraline (S; n = 42); 2) sertraline plus supervised group non-progressive exercise (S+NPE; n=37), or 3) sertraline plus supervised group progressive aerobic exercise (S+PAE; n=42). All patients received sertraline at a starting dosage of 50 mg, with later increases according to the clinical course. Participants in the S+NPE arm additionally attended three supervised group non-progressive exercise sessions (NP) per week in groups of 3-6 participants (60-minute duration). Patients in the S+PAE group attended

exercise sessions with a similar schedule to that of NP but exercised on bikes with a preplanned increase of the workload over the course of the study. The protocol also included brief sessions of interval training. Exercise sessions were supervised by medical and sport-science staff. Heart rate was continuously monitored to adapt the exercise intensity to individual aerobic capacity, which had been assessed by a peak oxygen uptake test. The total duration of the study was 24 weeks. The protocol of study interventions are briefly described in the supplementary materials (Appendix).

The primary outcome of the study was remission from depression, defined as a total score of 10 or less on the HAM-D at study end. A total of 15 participants withdrew from the study, but where included in the Intention to Treat analyses (Belvederi Murri et al., 2015). Six withdrew from the S group (four unwilling to continue, two for medical problems) and nine from the exercise groups (six unwilling to continue, two for medical problems and one for need of higher level of care). Since the groups receiving sertraline plus exercise displayed similar rates of remission from depression (S+PAE: 81%; S+NPE: 73%), for the aims of the present study they were combined into a sertraline plus exercise group (S+EX, n=79, remission rate: 77%), to be contrasted with the sertraline-only group (S, n=42, remission rate: 45%).

2.2 Assessment of symptoms

Depressive symptoms were assessed with the HAM-D at baseline, 4, 8, 12, and 24 weeks. Raters were certified psychiatrists experienced in psychogeriatrics. To improve inter-rater reliability, raters from each center participated in training sessions that included discussion of example cases. Symptom dimension scores were computed based on a previous factor analysis of the HAM-D conducted with a sample of 206 community-dwelling elderly individuals (Onega and Abraham, 1997). The analysis yielded four factors: (1) *affective* (depressed mood, guilt, suicide, work and activities, psychomotor retardation, loss of energy, loss of libido; items 1, 2, 3, 7, 8, 13, 14); (2) *vegetative* (initial, middle, and delayed insomnia, loss of appetite, loss of weight; items 4, 5, 6, 12, 16); (3) *anxiety* (psychological anxiety, somatic anxiety, hypochondriasis; items 10, 11, 15); (4) *agitation/insight* (agitation, lack of insight; items 9 and 17). To compute symptom dimension scores, item scores were summed and divided by the number of items in each factor. For descriptive purposes, at baseline, the participants also completed a battery of instruments assessing cognitive status, disability (Montreal Cognitive Assessment, MOCA) (Santangelo et al., 2014), physical comorbidities (Cumulative Illness Rating Scale, CIRS) (Miller et al., 1992) and other relevant variables (Neviani et al., 2017).

2.5 Statistical analysis

First, baseline characteristics of the two groups, including symptom dimensions scores, were compared by means of Chi-Square and T-tests, using an alpha level of 0.05. Second, to examine the impact of adjunctive exercise on the severity of symptom dimensions, we used multilevel Growth Curve Analysis (GCA). This analysis accounts for non-linear patterns of change over time (Shek and Ma, 2011). Symptom dimension scores at baseline, 4, 8, 12 and 24 weeks, nested within individuals, were the dependent variables. The trajectories of change were modeled as linear, quadratic, and cubic trends. Time and its interaction with group (S vs.S+EX) were treated as fixed effects. To account for potential confounders, the models were adjusted for physical comorbidities (CIRS severity index), sertraline dosage, age, and gender, after centering of the variables. The parameters were estimated using the Maximum Likelihood Method and Unstructured Covariance Structure. Alpha was adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate method.

Multilevel GCA models allow for missing data (Shek and Ma, 2011). Nonetheless, to verify the robustness of the results, we performed sensitivity analyses (for methods and results of sensitivity analyses, see the Supplementary Materials).

Third, we conducted exploratory analyses to examine the changes of individual symptoms scores (using each HAM-D item score as the dependent variable). For these exploratory analyses, the alpha level was kept at 0.05. Effect sizes were calculated to quantify the magnitude and timing of the effects on symptom dimensions and the severity of individual symptoms. Effect sizes were computed as the difference between the effect in the experimental group (S+EX) minus that in the control group (S) at each time point, divided by the baseline standard deviation of the respective variable (Feingold, 2009). For ease of interpretation, a positive value represents greater improvement (reduced severity of the symptom dimension) in the S+EX group relative to the S group. Analyses were performed using SPSS version 17.0.

3. Results

3.1 Recruitment and participant characteristics

Baseline characteristics of participants are shown in Table 1. At baseline, there were no inter-group differences for most sociodemographic and clinical characteristics, except for a higher burden of physical comorbidities in the S+EX group. Likewise, the levels of disability and the severity of symptom dimensions did not differ between the S and S+EX groups, and there were no significant inter-group differences in the severity of individual symptoms endorsed by participants (all $p>0.10$ for comparisons of frequency and severity), with the exception of delayed insomnia (S: 1.24±0.85, S+EX: 0.87±0.84, $p=0.03$). Consistent with the orthogonal rotation in the original factor analysis,

symptom dimension scores exhibited non-significant inter-correlations, with the exception of the correlation between depressed affect and agitation/insight ($r= -0.38$, $p<0.001$; see Supplementary Materials, Table S5).

3.2 Changes in the severity of symptom dimensions

By study end, patients receiving the adjunctive exercise interventions (S+EX) showed lower unadjusted scores for affective and vegetative symptom dimensions (both $p<0.025$, corrected for Benjamini-Hochberg False Discovery Rate) compared to patients on sertraline. On the other hand, there was no significant difference in the severity of agitation/insight ($p=0.55$) or anxiety ($p=0.06$). Results from Growth Curve Analyses showed that all participants displayed significant time-dependent reductions in the four symptom dimensions compared to baseline (see Table 2, left panel). However, the S+EX group displayed larger reductions in the *affective* symptom dimension than the S group (Figure 1), although there were no significant between-group differences in the trajectories of other symptom dimensions (all $p>0.020$, corrected for Benjamini-Hochberg False Discovery Rate). Table 3 shows the corresponding effect sizes: a large effect size ($d=0.79$) indicates improvement of affective symptoms over the entire study period, with greater effects in the first 4 weeks ($d=0.54$) and during the second half of the study ($d=0.72$, from 12 to 24 weeks) in the S+EX group compared to S. Sensitivity analyses revealed that the results were robust to missing data (see Supplementary Materials, Table S6).

3.3 Changes in the severity of individual symptoms

Exploratory GCAs were conducted to examine changes of individual depressive symptoms (see Supplementary Materials, Table S5). Models for items 3, 4, 5, and 6 (suicide and three insomnia items) did not reach convergence and, therefore, are not reported. All other symptoms, except for insight and retardation, showed time-dependent changes indicating decreased severity from baseline in the whole sample. Moreover, those in the S+EX group, relative to the S group, displayed significantly greater improvements in depressed mood (item 1), psychomotor retardation (item 8), and psychic anxiety (item 10), as evident from significant group x time (linear and/or quadratic and/or cubic terms) interactions (see Supplementary Materials, Figures S5 – S7). Effect sizes, in the small-to-medium range, are reported in Table 3. Other symptoms (guilt, work and activities, agitation, somatic anxiety, general and gastrointestinal somatic symptoms, genital symptoms, hypochondriasis, and weight loss) did not show significant between-group changes (all $p>0.10$).

4. Discussion

This study examined the timing and profile of the clinical response to exercise in a sample of elderly patients with major depression. Compared with individuals receiving only standard antidepressant treatment, those who were additionally treated with exercise displayed greater improvements in the affective symptom dimension, especially in the first 4 weeks of treatment and after 12 weeks. The study recruited a representative sample of older primary-care depressed individuals, who tend to be characterized by suboptimal response to antidepressant drugs and frequent residual symptoms (Alexopoulos, 2005).

Evidence supporting the effectiveness of exercise against late-life depression has proliferated over the last several years (Heinzel et al., 2015; Schuch et al., 2016b; Vancampfort et al., 2017). However, it is still partly unclear whether exercise improves mainly the somatic symptoms of depression, which may be non-specific indicators of disease, or can also address "core" depressive symptoms. Most prior studies have not examined this issue in detail, since outcome reporting has typically focused on total scores of rating scales (Schuch et al., 2016b). To our knowledge, only three studies have previously examined symptom dimensions among the elderly (Lavretsky et al., 2011; Singh et al., 1997; Singh et al., 2001; Singh et al., 2005), and in only one of these studies was antidepressant drug treatment included as a comparator (Lavretsky et al., 2011). In that study, older depressed individuals who had not remitted after treatment with escitalopram were subsequently randomized to receive additional Tai-chi or health education meetings. Those who attended this light-intensity exercise displayed higher remission rates and greater improvements in symptoms of apathy, though not anxiety (Lavretsky *et al.*, 2011). In another study of older adults, low- and high-intensity exercise were compared with "unrestricted" treatment of depression by primary-care physicians. Exercise was associated with greater improvements in sleep quality and the overall severity of depression (Singh *et al.*, 2005). However, the study did not specifically examine mood or other core symptoms of depression, included individuals with both major and minor depression, and only some of the participants received antidepressant drug treatment (42% of those in the control group, none in the exercise groups). Finally, Singh and colleagues randomized depressed older adults to either a supervised progressive resistance training program or to an attention-control group (Singh *et al.*, 1997;Singh *et al.*, 2001). Although the effects were stronger on somatic than "psychological" symptoms, the group by time interactions for the "psychological" symptoms of the Beck Depression Inventory fell just short of statistical significance, due to the low statistical power afforded by the small sample.

In our study, participants receiving both antidepressants and exercise showed larger changes in affective symptoms, likely driven by improvements in mood and psychomotor retardation. Exercise is known to induce pleasurable affective states, provided that its intensity is either self-regulated or tailored to the individual fitness level (Ekkekakis et al., 2011); indeed, in our study, exercise intensity was carefully tailored to individual physical fitness level and monitored by personnel throughout the sessions to avoid overexertion and, therefore, unpleasant affective responses. A positive affective experience from early sessions has been shown to favor subsequent patient compliance and ultimately increase the antidepressant efficacy of exercise (Suterwala et al., 2016). At the biological level, exercise may acutely impact on mood and psychomotor retardation through several mechanisms, including the modulation of monoaminergic neurotransmission, neural circuits involved in somatosensory and mood regulation, and via adaptations in hypothalamic-pituitary-adrenal axis activity (Buyukdura et al., 2011; Ekkekakis, 2013; Tozzi et al., 2016). Moreover, in our study, adding exercise to antidepressants was associated with improvements in autonomic nervous system balance (Toni et al., 2016) and cognitive function (Neviani et al., 2017), which may have contributed to changes in affective and motor symptoms. It is also possible that different mechanism underlie changes in each symptom dimension; for instance exercise may improve psychomotor retardation directly, influencing physical condition (Henderson et al., 2017) or indirectly, by improving cognitive functions that partly underlie this symptom in older individuals (Buyukdura et al., 2011; Gabel et al., 2015). Notably, studies that specifically examine the biological responses to exercise among patients with late-life depression are still greatly needed (Schuch et al., 2016a).

Another important issue is that of timing. In our study, exercise started within days after the initiation of antidepressant drug treatment and was associated with both early (first 4 weeks of treatment) and later (after 12 weeks) advantages over drug therapy alone. On the other hand, we observed a relative "plateau" in the middle (4 to 12 weeks), where adjunctive exercise evidently did not confer additional benefits. Biphasic improvements have also been observed among younger adults (Legrand et al., 2009; Legrand and Neff, 2016) and may depend on the onset of slower mechanisms, such as neurogenesis (Kerling et al., 2017; Thomas et al., 2012). Alternatively, antidepressant response among elderly patients requires up to 4-6 weeks to become evident (although not all studies are concordant; (Whyte et al., 2004), so it is possible that the effects of sertraline emerged during this phase, thus reducing the differences with adjunctive exercise.

The strengths of this study include a representative sample of old individuals (mean age 75) diagnosed with major depression, a long-term exercise intervention with repeated assessments, and a robust analytic approach. However, the findings reported here must be viewed in light of certain limitations. First, the sample size was relatively small, making it necessary to combine participants from the

"aerobic progressive" and "non-progressive" groups and possibly averaging out differential effects due to these different types of exercise. Second, results pertaining to single-item analyses were not adjusted for multiple comparisons, thus they must be viewed with caution. Nonetheless, these results show that most of the efficacy of exercise was observed on affective, rather than somatic, symptoms or anxiety. Third, despite randomization, the groups displayed baseline differences in the severity of physical comorbidities and, at study end, participants in the S group were receiving slightly higher doses of sertraline than those in the S+EX group. It should be kept in mind, however, that analyses were adjusted for these confounders. Fourth, depressive symptoms were rated only with the HAM-D, while the use of other or additional instruments might have allowed a more fine-grained assessment. Fifth, given the absence of an exercise-only arm, we are not able to make inferences regarding possible interactions (e.g., synergistic effects) between sertraline and exercise. Finally, given the absence of a placebo arm, it is not possible to disentangle the effects of sertraline or exercise from those due to expectancy.

In conclusion, compared with standard antidepressant treatment, exercise plus sertraline was associated with additional improvements in the affective domain of depression, including mood and psychomotor retardation. Thus, even among the elderly, the benefit of adding exercise to drug treatment does not appear to reply on the amelioration of somatic symptoms or anxiety, but rather on nuclear symptoms of depression. Clinicians treating late-life depression should be aware that these symptoms may improve both in the short-term and after months of treatment.

Alexopoulos, G.S., 2005. Depression in the elderly. Lancet 365, 1961-1970.

Belvederi Murri, M., Amore, M., Menchetti, M., Toni, G., Neviani, F., Cerri, M., Rocchi, M.B., Zocchi, D., Bagnoli, L., Tam, E., Buffa, A., Ferrara, S., Neri, M., Alexopoulos, G.S., Zanetidou, S., 2015. Physical exercise for late-life major depression. Br J Psychiatry 207, 235-242.

Buyukdura, J.S., McClintock, S.M., Croarkin, P.E., 2011. Psychomotor retardation in depression: biological underpinnings, measurement, and treatment. Progress in neuro-psychopharmacology & biological psychiatry 35, 395-409.

Ekkekakis, P., 2013. Routledge handbook of physical activity and mental health. Routledge.

Ekkekakis, P., 2015. Honey, I shrunk the pooled SMD! Guide to critical appraisal of systematic reviews and meta-analyses using the Cochrane review on exercise for depression as example. Mental Health and Physical Activity 8, 21-36.

Ekkekakis, P., Belvederi Murri, M., 2017. Exercise as antidepressant treatment: Time for the transition from trials to clinic? Elsevier.

Ekkekakis, P., Parfitt, G., Petruzzello, S.J., 2011. The pleasure and displeasure people feel when they exercise at different intensities: decennial update and progress towards a tripartite rationale for exercise intensity prescription. Sports Med 41, 641-671.

Feingold, A., 2009. Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. Psychol Methods 14, 43-53.

Gabel, N.M., Crane, N.A., Avery, E.T., Kay, R.E., Laurent, A., Giordani, B., Alexander, N.B., Weisenbach, S.L., 2015. Dual-tasking gait variability and cognition in late-life depression. Int J Geriatr Psychiatry 30, 1120-1128.

Groeneweg-Koolhoven, I., Ploeg, M., Comijs, H.C., Wjh Penninx, B., van der Mast, R.C., Schoevers, R.A., Rhebergen, D., Exel, E.V., 2017. Apathy in early and late-life depression. Journal of affective disorders 223, 76-81.

Haigh, E.A.P., Bogucki, O.E., Sigmon, S.T., Blazer, D.G., 2017. Depression Among Older Adults: A 20-Year Update on Five Common Myths and Misconceptions. Am J Geriatr Psychiatry.

Hegeman, J.M., de Waal, M.W., Comijs, H.C., Kok, R.M., van der Mast, R.C., 2015. Depression in later life: a more somatic presentation? J Affect.Disord 170, 196-202.

Heinzel, S., Lawrence, J.B., Kallies, G., Rapp, M.A., Heissel, A., 2015. Using exercise to fight depression in older adults: A systematic review and meta-analysis. Hogrefe Publishing.

Henderson, R.M., Leng, X.I., Chmelo, E.A., Brinkley, T.E., Lyles, M.F., Marsh, A.P., Nicklas, B.J., 2017. Gait speed response to aerobic versus resistance exercise training in older adults. Aging clinical and experimental research 29, 969-976.

Iniesta, R., Malki, K., Maier, W., Rietschel, M., Mors, O., Hauser, J., Henigsberg, N., Dernovsek, M.Z., Souery, D., Stahl, D., Dobson, R., Aitchison, K.J., Farmer, A., Lewis, C.M., McGuffin, P., Uher, R., 2016. Combining clinical variables to optimize prediction of antidepressant treatment outcomes. J Psychiatr Res 78, 94-102.

Kerling, A., Kuck, M., Tegtbur, U., Grams, L., Weber-Spickschen, S., Hanke, A., Stubbs, B., Kahl, K.G., 2017. Exercise increases serum brain-derived neurotrophic factor in patients with major depressive disorder. Journal of affective disorders 215, 152-155.

Lavretsky, H., Alstein, L.L., Olmstead, R.E., Ercoli, L.M., Riparetti-Brown, M., Cyr, N.S., Irwin, M.R., 2011. Complementary use of tai chi chih augments escitalopram treatment of geriatric depression: a randomized controlled trial. Am J Geriatr Psychiatry 19, 839-850.

Legrand, F.D., Bertucci, W.M., Thatcher, J., 2009. Telic dominance influences affective response to a heavy-intensity 10-min treadmill running session. J Sports Sci. 27, 1059-1067.

Legrand, F.D., Neff, E.M., 2016. Efficacy of exercise as an adjunct treatment for clinically depressed inpatients during the initial stages of antidepressant pharmacotherapy: An open randomized controlled trial. Journal of affective disorders 191, 139-144.

Luck-Sikorski, C., Stein, J., Heilmann, K., Maier, W., Kaduszkiewicz, H., Scherer, M., Weyerer, S., Werle, J., Wiese, B., Moor, L., Bock, J.O., Konig, H.H., Riedel-Heller, S.G., 2017. Treatment preferences for depression in the elderly. International psychogeriatrics 29, 389-398.

Miller, M.D., Paradis, C.F., Houck, P.R., Mazumdar, S., Stack, J.A., Rifai, A.H., Mulsant, B., Reynolds, C.F., III, 1992. Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. Psychiatry Res 41, 237-248.

Naismith, S.L., Norrie, L.M., Mowszowski, L., Hickie, I.B., 2012. The neurobiology of depression in later-life: clinical, neuropsychological, neuroimaging and pathophysiological features. Prog.Neurobiol. 98, 99-143.

Nelson, M.E., Rejeski, W.J., Blair, S.N., Duncan, P.W., Judge, J.O., King, A.C., Macera, C.A., Castaneda-Sceppa, C., 2007. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. Circulation 116, 1094-1105.

Neviani, F., Belvederi Murri, M., Mussi, C., Triolo, F., Toni, G., Simoncini, E., Tripi, F., Menchetti, M., Ferrari, S., Ceresini, G., Cremonini, A., Bertolotti, M., Neri, G., Squatrito, S., Amore, M.,

Zanetidou, S., Neri, M., 2017. Physical exercise for late life depression: effects on cognition and disability. International psychogeriatrics 29, 1105-1112.

Onega, L.L., Abraham, I.L., 1997. Factor structure of the Hamilton Rating Scale for Depression in a cohort of community-dwelling elderly. Int J Geriatr Psychiatry 12, 760-764.

Santangelo, G., Siciliano, M., Pedone, R., Vitale, C., Falco, F., Bisogno, R., Siano, P., Barone, P., Grossi, D., Santangelo, F., Trojano, L., 2014. Normative data for the Montreal Cognitive Assessment in an Italian population sample. Neurol.Sci.

Schuch, F.B., Deslandes, A.C., Stubbs, B., Gosmann, N.P., Silva, C.T., Fleck, M.P., 2016a. Neurobiological effects of exercise on major depressive disorder: A systematic review. Neurosci.Biobehav.Rev 61, 1-11.

Schuch, F.B., Vancampfort, D., Rosenbaum, S., Richards, J., Ward, P.B., Veronese, N., Solmi, M., Cadore, E.L., Stubbs, B., 2016b. Exercise for depression in older adults: a meta-analysis of randomized controlled trials adjusting for publication bias. Revista brasileira de psiquiatria 38, 247-254.

Shek, D.T., Ma, C.M., 2011. Longitudinal data analyses using linear mixed models in SPSS: concepts, procedures and illustrations. ScientificWorldJournal. 11, 42-76.

Singh, N.A., Clements, K.M., Fiatarone, M.A., 1997. A randomized controlled trial of progressive resistance training in depressed elders. J Gerontol.A Biol Sci.Med Sci. 52, M27-M35.

Singh, N.A., Clements, K.M., Singh, M.A., 2001. The efficacy of exercise as a long-term antidepressant in elderly subjects: a randomized, controlled trial. J Gerontol.A Biol Sci.Med Sci. 56, M497-M504.

Singh, N.A., Stavrinos, T.M., Scarbek, Y., Galambos, G., Liber, C., Fiatarone Singh, M.A., 2005. A randomized controlled trial of high versus low intensity weight training versus general practitioner care for clinical depression in older adults. J Gerontol.A Biol Sci.Med Sci. 60, 768-776.

Stubbs, B., Rosenbaum, S., Vancampfort, D., Ward, P.B., Schuch, F.B., 2016a. Exercise improves cardiorespiratory fitness in people with depression: A meta-analysis of randomized control trials. J Affect.Disord 190, 249-253.

Stubbs, B., Vancampfort, D., Rosenbaum, S., Ward, P.B., Richards, J., Soundy, A., Veronese, N., Solmi, M., Schuch, F.B., 2016b. Dropout from exercise randomized controlled trials among people with depression: A meta-analysis and meta regression. J Affect.Disord 190, 457-466.

Suterwala, A.M., Rethorst, C.D., Carmody, T.J., Greer, T.L., Grannemann, B.D., Jha, M., Trivedi, M.H., 2016. Affect Following First Exercise Session as a Predictor of Treatment Response in Depression. J Clin Psychiatry 77, 1036-1042.

Thomas, A.G., Dennis, A., Bandettini, P.A., Johansen-Berg, H., 2012. The effects of aerobic activity on brain structure. Front Psychol 3, 86.

Toni, G., Belvederi, M.M., Piepoli, M., Zanetidou, S., Cabassi, A., Squatrito, S., Bagnoli, L., Piras, A., Mussi, C., Senaldi, R., Menchetti, M., Zocchi, D., Ermini, G., Ceresini, G., Tripi, F., Rucci, P., Alexopoulos, G.S., Amore, M., 2016. Physical Exercise for Late-Life Depression: Effects on Heart Rate Variability. Am J Geriatr Psychiatry 24, 989-997.

Tozzi, L., Carballedo, A., Lavelle, G., Doolin, K., Doyle, M., Amico, F., McCarthy, H., Gormley, J., Lord, A., O'Keane, V., Frodl, T., 2016. Longitudinal functional connectivity changes correlate with mood improvement after regular exercise in a dose-dependent fashion. The European journal of neuroscience 43, 1089-1096.

Uher, R., Maier, W., Hauser, J., Marusic, A., Schmael, C., Mors, O., Henigsberg, N., Souery, D., Placentino, A., Rietschel, M., Zobel, A., Dmitrzak-Weglarz, M., Petrovic, A., Jorgensen, L., Kalember, P., Giovannini, C., Barreto, M., Elkin, A., Landau, S., Farmer, A., Aitchison, K.J.,

McGuffin, P., 2009. Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression. Br J Psychiatry 194, 252-259.

Uher, R., Perlis, R.H., Henigsberg, N., Zobel, A., Rietschel, M., Mors, O., Hauser, J., Dernovsek, M.Z., Souery, D., Bajs, M., Maier, W., Aitchison, K.J., Farmer, A., McGuffin, P., 2012. Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. Psychol Med 42, 967-980.

Vancampfort, D., Hallgren, M., Firth, J., Rosenbaum, S., Schuch, F.B., Mugisha, J., Probst, M., Van Damme, T., Carvalho, A.F., Stubbs, B., 2017. Physical activity and suicidal ideation: A systematic review and meta-analysis. Journal of affective disorders 225, 438-448.

Whyte, E.M., Dew, M.A., Gildengers, A., Lenze, E.J., Bharucha, A., Mulsant, B.H., Reynolds, C.F., 2004. Time course of response to antidepressants in late-life major depression: therapeutic implications. Drugs Aging 21, 531-554.

Table 1. Baseline characteristics of participants

	S (n=42)	S+EX (n=79)	Statistics
Sociodemographic			
Age, mean (SD)	75.6 (5.6)	74.9 (6.2)	t=0.39, p=0.53
Gender, F (%)	76.2	68.4	χ^2=0.82, p=0.37
Marital status, single (%)	54.8	54.4	χ^2=0.01, p=0.97
Education, elementary or less (%)	64.3	48.1	χ^2=2.89, p=0.09
Living alone (%)	45.2	44.3	χ^2=0.01, p=0.92
Physical-medical			
BMI, mean (SD)	25.8 (3.3)	26.0 (3.8)	t=0.07, p=0.79
CIRS severity index, mean (SD)	1.31 (0.22)	1.40 (0.23)	t=4.10, p=0.05 *
CIRS comorbidity index, mean (SD)	0.57 (0.80)	1.26 (1.16)	t=12.9, p=0.001 *
Peak VO_2, mean (SD)	15.8 (2.7)	15.3 (3.6)	t=0.32, p=0.58
MOCA total score, mean (SD)	21.4 (4.2)	21.6 (4.1)	t=0.03, p=0.86
Psychiatric-cognitive			
Brief Disability Questionnaire, mean (SD)	9.6 (4.3)	9.5 (4.5)	t=0.16, p=0.87
Onset of depression after 55 years, %	46.3	45.5	χ^2=0.01, p=0.93
Treated with antidepressants lifetime (%)	73.8	64.6	χ^2=1.08, p=0.30
>2 depressive episodes lifetime (%)	42.9	31.0	χ^2=1.45, p=0.23
History of suicide attempt (%)	2.4	2.7	χ^2=0.01, p=0.93
HAM-D total score, mean (SD)	20.4 (3.4)	20.0 (2.9)	t=0.66, p=0.42
Symptom dimension scores (HAM-D factors)			
Baseline			
Affective	1.25 (0.32)	1.25 (0.35)	t=0.01, p=0.99
Vegetative	0.96 (0.37)	0.85 (0.41)	t=1.37, p=0.17
Anxiety	1.75 (0.72)	1.73 (0.65)	t=0.16, p=0.88
Agitation/Insight	0.82 (0.56)	0.87 (0.56)	t=0.49, p=0.63
Study end			
Affective	0.63 (0.42)	0.43 (0.35)	t=2.36, p=0.02 *
Vegetative	0.46 (0.38)	0.28 (0.30)	t=2.63, p=0.01 *
Anxiety	1.00 (0.67)	0.74 (0.66)	t=1.85, p=0.06
Agitation/insight	0.52 (0.44)	0.46 (0.48)	t=0.60, p=0.55

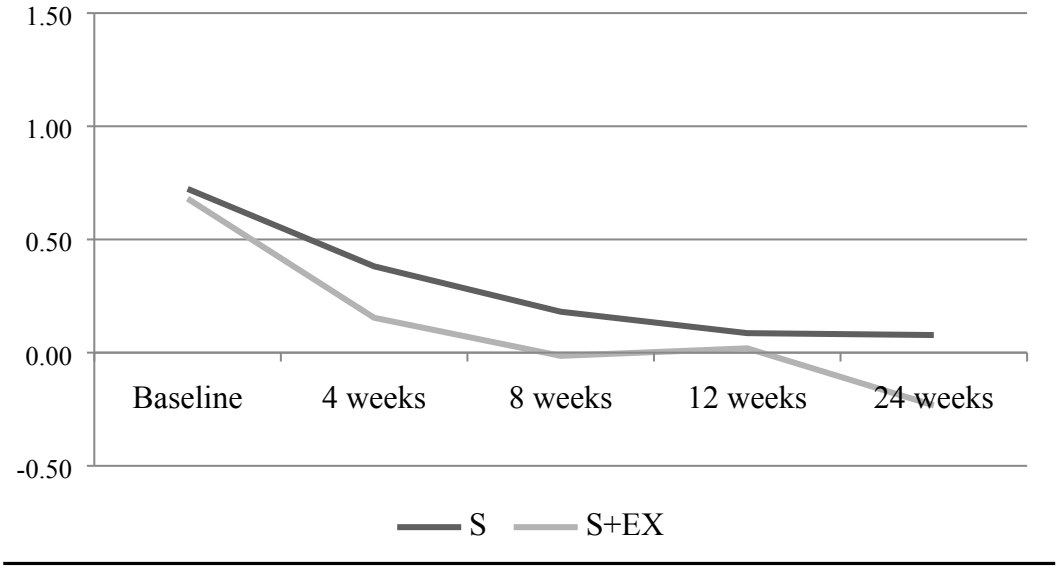
S, sertraline; S+EX, sertraline plus physical exercise * $p<0.05$

Table 2. Effect of sertraline plus exercise vs. sertraline alone on symptom dimensions

		Effect of time					Effect of group (S+EX vs. S)			
Symptom dimensions		parameter	Std. Error	df	p		parameter	Std. Error	df	p
Affective	intercept	0.709206	0.336986	127.9881	0,04	S/EX	-0.01806	0.032277	200.8038	0.58
	time	-0.11927	0.012552	339.3405	<0.001 *	S+EX x time	-0.03861	0.012506	337.7333	0.002 *
	time^2	0.006944	0.001562	323.5803	<0.001 *	S+EX x time^2	0.004497	0.001563	323.7473	0.004 *
	time^3	-0.00014	0.000045	322.642	0.003 *	S+EX x time^3	-0.00013	0.000454	322.8255	0.006 *
Vegetative	intercept	0.2977	0.337381	126.5232	0.38	S/EX	-0.06339	0.036531	176.3947	0.08
	time	-0.11828	0.012682	338.3671	<0.001 *	S+EX x time	-0.00041	0.012639	336.683	0.97
	time^2	0.008021	0.001583	324.5043	<0.001 *	S+EX x time^2	0.000526	0.001582	324.6544	0.74
	time^3	-0.00017	0.000046	323.6541	<0.001 *	S+EX x time^3	-0.00002	0.000046	323.8491	0.60
Anxiety	intercept	1.942719	0.657968	127.5344	0.004 *	S+EX	-0.03882	0.062953	179.7882	0.54
	time	-0.14065	0.021382	344.066	<0.001 *	S+EX x time	-0.03875	0.021303	342.5337	0.07
	time^2	0.007343	0.002656	328.4478	0.006 *	S+EX x time^2	0.003454	0.002657	328.612	0.19
	time^3	-0.00012	0.000077	327.7846	0.11	S+EX x time^3	-0.000086	0.000077	327.9621	0.28
Agitation/Insight	intercept	1.467181	0.486882	130.4629	0.003 *	S+EX	0.032075	0.050118	196.635	0.52
	time	-0.05579	0.019237	341.2208	0.004 *	S+EX x time	-0.03029	0.019172	339.4689	0.12
	time^2	0.003064	0.002396	326.4069	0.20	S+EX x time^2	0.003135	0.002397	326.5751	0.19
	time^3	-0.00005	0.000071	325.4669	0.39	S+EX x time^3	-0.00000005	0.000071	325.6689	0.25

Analyses are based on hierarchical Growth Curve Analyses in the completers dataset. Symptom dimension scores, nested within individuals, are used as the dependent variables. The linear, quadratic and cubic term of time are modeled together with their interaction with group (S+EX vs. S) and adjusted for age, gender, CIRS severity index, and changes in sertraline dosages. The significance levels are corrected for multiple testing using Benjamini and Hochberg false discovery rate ($p<0.02$). A negative parameter indicates greater decrease of symptom dimension severity in the experimental condition (S+EX) relative to the control condition (S). For example, a negative parameter for "S+EX x time^2" indicates that there is a greater quadratic decline in the S+EX group compared with the S group.

Figure 1. Changes in Affective symptom dimension scores



The graph is based on model estimate scores after adjustment for covariates

Table 3. Effects sizes of sertraline plus exercise on symptom dimensions and individual symptoms

		0 - 4 weeks	4 - 8 weeks	8 - 12 weeks	12 - 24 weeks	0 -24 weeks
Symptom dimensions (HAM-D factors)						
Affective		0.54	-0.08	-0.35	0.72	0.79
Vegetative		-0.21	-0.05	-0.07	0.38	0.09
Anxiety		0.34	0.05	-0.09	0.17	0.46
Agitation/Insight		0.28	-0.01	-0.17	0.07	0.18
Symptoms (HAM-D items)						
item 1	Depressed mood	0.38	-0.10	-0.28	0.27	0.19
item 8	Retardation	0.50	-0.09	-0.50	0.41	0.33
item 10	Anxiety - psychic	0.39	0.09	-0.08	0.04	0.45

The table reports effect sizes associated with the experimental intervention (S+EX) relative to the comparator (S), during each time interval. A positive effect size indicates a greater improvement of symptom dimension score (HAM-D factors) in the S+EX group relative to the S group.

Conflicts of interest: none

Funding. This research was supported by the Italian Emilia Romagna Region University Programme (PrRU) 2010–2012 grant, Area 2 for Clinical Governance. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. All authors declare that they have no conflicts of interest

Contributors. M. Belvederi Murri designed the study, collected clinical data, performed statistical analyses and wrote the article. P. Ekkekakis supervised statistical analyses and wrote the article. M. Neri, M. Amore, S. Zanetidou, M. Menchetti, F. Neviani, S. Squatrito, F. Trevisani and G. Toni, designed the study, participated in data collection and wrote the article. S. Tedeschi, E. Nerozzi, P. Maietta Latessa, G. Ermini and D. Zocchi participated in data collection and assisted with writing the article.

Acknowledgements. Other members of the Safety and Efficacy of Exercise for Depression in Seniors (SEEDS) Study Group were involved at various titles in the study design, clinical assessment of patients, organization, management and delivery of the study interventions. **Bologna Centers:** Francesco Ripa di Meana, MD, Maria L. Marcaccio, MD, Paola Argnani, MD, Angelo Fioritti, MD, Mara Morini, MD, Roberto Senaldi, MD, Giovanni De Girolamo, MD, PhD, Danilo Di Diodoro, MD, Giovanni Neri, MD, Matteo Cerri MD, PhD, Marco BL Rocchi PhD, Serena Ferrara PsyD, Piero Casarini, MD, Luigi Bagnoli MD, Carlo Spezia, MD, Monica Magagnoli, MD, Claudia Luciano, MD, Federica Casini, PsyD, Angela Buffa MD, Pier Vittorio Bardazzi, MD, Alessandro Piras, PhD, Giuliana Tola, Roberto Rizzo, Lorenzo De Bei, Davide Comastri. **Modena-Correggio Center**: Ferdinando Tripi, MD, Massimo Mannina, MD; Massimo Piepoli, MD, PhD, Sandro Zoboli, Gabriele Torcianti, Niccolò Colombini, MD, PhD, Guido Federzoni, MD, Morena Pellati, MD, Veronica Barbanti Silva, MD, Agnese Cepelli, Elena Francia PsyD, Roberto Chiesa, MD, Anna Grazia Frignani, MD. **Parma Centre**: Cristina Montegano, MD, PhD, Arianna Montali, MD, Aderville Cabassi, MD, PhD, Mattia Masotti, MD, PhD, Maria Lidia Gerra, MD. The authors would also like to thank Prof. George S. Alexopoulos for his kind mentoring and precious support. The authors wish also to thank participants who took part in the study.

Supplementary material for online publication

Supplementary methods

To verify the robustness of results relative to the changes of symptom dimensions, we performed sensitivity analyses. Growth Curve Analyses were repeated after imputation of missing data with two methods. First, we employed the Last-Observation-Carried Forward (LOCF) method. LOCF assumes non-random distribution of missing values and that the scores of participants who drop-out from the study remain equal to the last observed response (Dziura et al., 2013). The second method was based on substituting missing data with Multiple Imputation (MI). MI works under the assumption that missing data are randomly distributed, and generates different sets of plausible values for the missing observations, which are combined in subsequent analyses (Dziura et al., 2013). Five regression-based MI sets were generated, using baseline depression severity score (HAM-D), age, and gender as predictors.

Supplementary results

Table S5. Correlation matrix between baseline symptom dimension scores

		Depressed affect	Anxiety	Vegetative symptoms	Agitation/ insight
Affective	R	1	-0.17	-0.11	-0.38 **
	p		0.06	0.24	<0.001
Anxiety	R		1	0.01	-0.15
	p			0.96	0.10
Vegetative	R			1	-0.02
	p				0.83
Agitation/insight	R				1
	p				

Changes in symptom dimensions

Figure S2. Changes in severity of the vegetative symptom dimension in the study groups

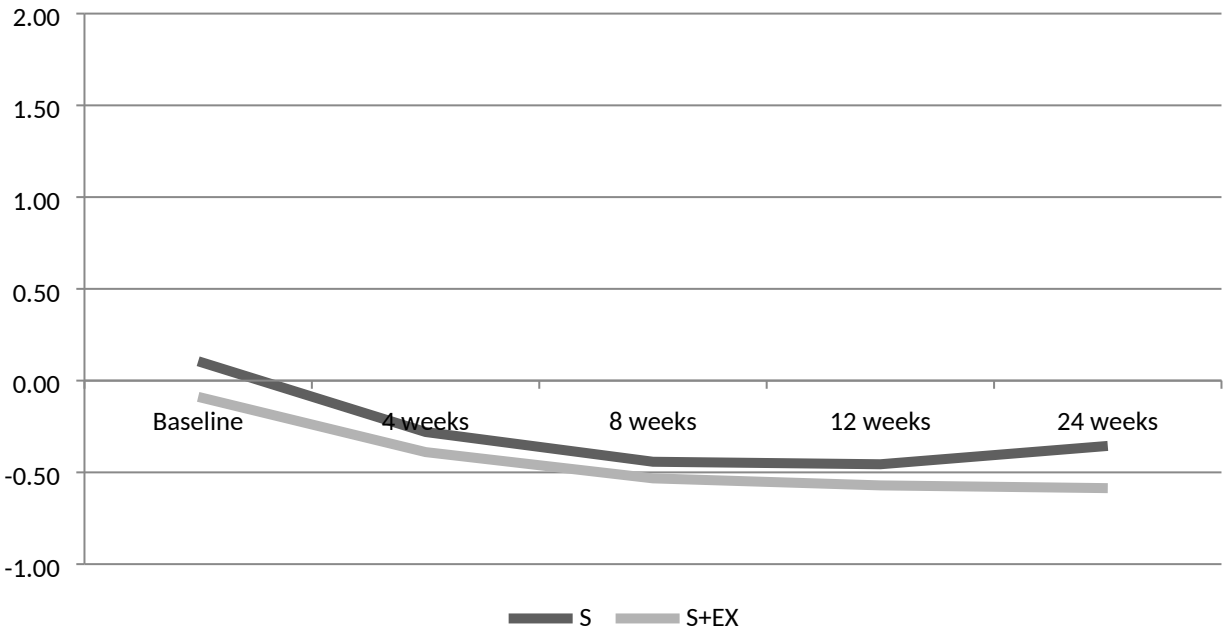


Figure S3. Changes in severity of the anxiety symptom dimension in the study groups

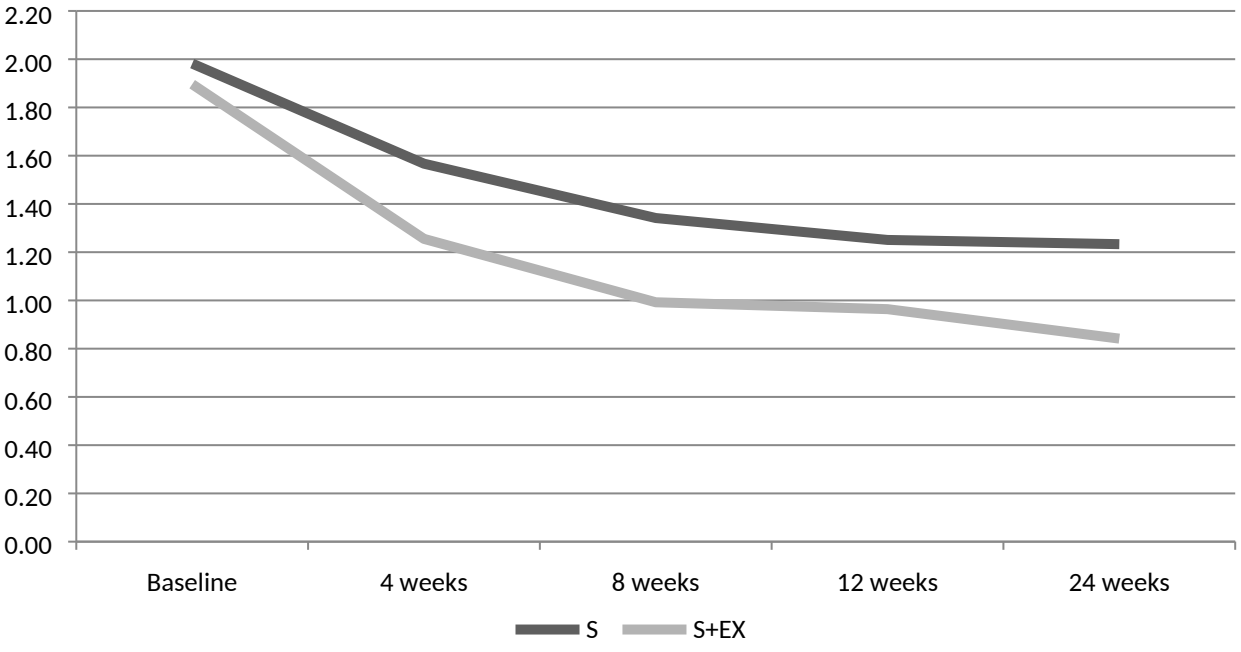
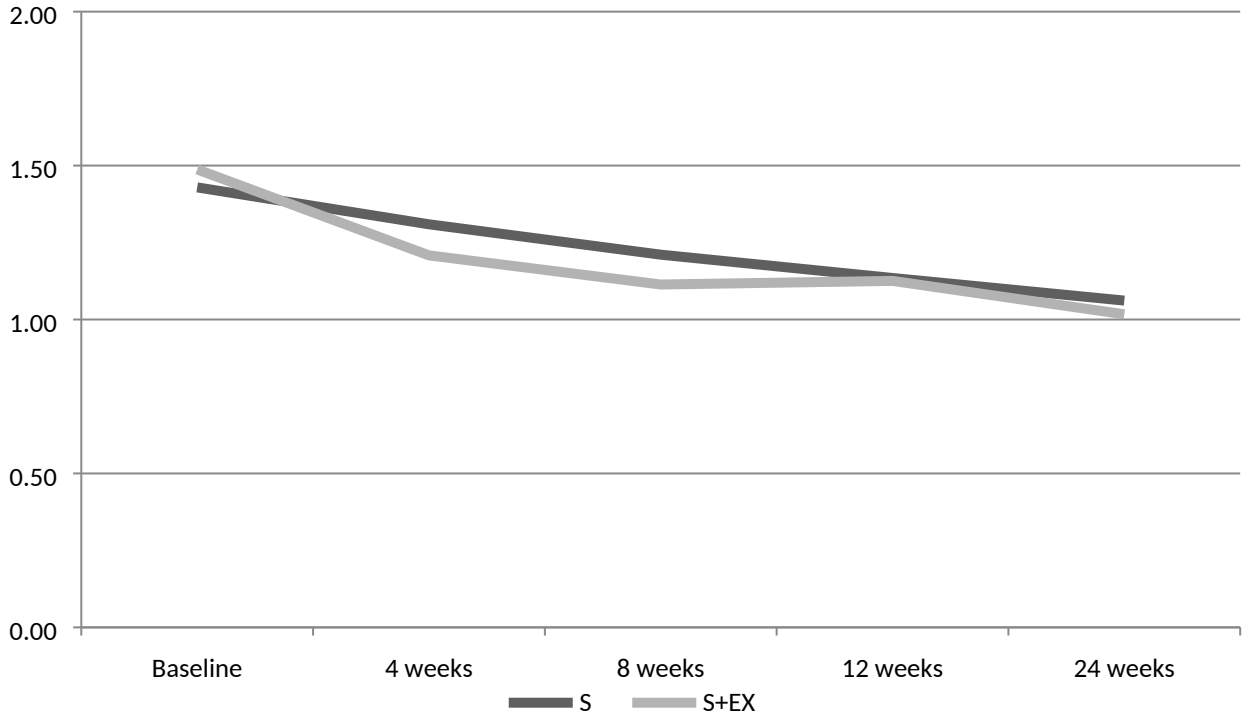


Figure S4. Changes in severity of the agitation/insight symptom dimension in the study groups



Sensitivity analyses

Table S6. Sensitivity analyses: effect sizes of the experimental intervention on symptom dimensions

	0 - 4 weeks	4 - 8 weeks	8 - 12 weeks	12 - 24 weeks	0 -24 weeks
Completers dataset					
Affective	0.54	-0.08	-0.35	0.72	0.79 *
Vegetative	-0.21	-0.05	-0.07	0.38	0.09
Anxiety	0.34	0.05	-0.09	0.17	0.46
Agitation/insight	0.28	-0.01	-0.17	0.07	0.18
LOCF imputation					
Affective	0.60	0.31	-0.22	0.24	0.65 *
Vegetative	0.01	-0.04	-0.05	0.39	0.25
Anxiety	0.37	0.12	-0.02	0.00	0.47
Agitation/insight	0.41	0.04	-0.13	-0.05	0.27
Multiple imputation					
Affective	0.57	-0.03	-0.30	0.38	0.60 *
Vegetative	0.01	-0.07	-0.09	0.46	0.21
Anxiety	0.31	0.07	-0.06	0.08	0.40
Agitation/insight	0.31	0.02	-0.15	-0.08	0.12

A positive effect size indicates a greater improvement of symptom dimension in the S+EX group relative to the S group. The upper panel reports effect sizes obtained from the completers dataset (with missing values), the lower panels from two imputed datasets. All effect sizes are calculated from model estimates. * result is statistically significant at $p<0.02$ after correction for multiple testing.

Analyses of the severity of symptom dimensions were repeated after imputation of missing data. Baseline values for symptom dimensions were complete, whereas the number of cases with missing data at subsequent time points ranged from 9 to 22 (7.4 – 18.2%). No clear pattern emerged from the examination of the distribution of missing data, and the Little test indicated that data could be considered missing completely at random (χ^2 =225.9, df= 212, p=0.25).

Analyses of the LOCF-imputed data did not yield substantially different results from those of the analyses bases on the sample of completers. Compared with the S group, those in the S+EX group had significantly larger improvements of the *affective* symptom dimension ($p<0.02$) but not of vegetative symptoms, anxiety, or agitation/insight (all $p>0.02$). The overall effect size of the intervention was medium, slightly smaller than that obtained from the sample of completers (0.65). Considering the timing of improvements, the larger effects on affective symptoms were evident in the first 8 weeks and in the second month of the study.

Analyses of the MI dataset were again similar to those obtained from the sample of completers. The S+EX group exhibited larger decreases of affective symptoms (all $p<0.02$) but not of vegetative symptoms, anxiety, or agitation/insight (all $p>0.02$). The overall effect size of the intervention was again slightly smaller than in the completers dataset (0.60). The timing of the effects on affective symptoms was mostly evident in the first 4 weeks and in the last 12 weeks of treatment.

Table S6 Changes in severity of symptoms

		Model					Effect of group (S+EX vs. S)			
Symptom dimensions		parameter	Std. Error	df	p		parameter	Std. Error	df	p
1. Depressed Mood	intercept	2.713	0.817	125.325	0.001 *	S/EX	-0.137	0.087	214.449	0.119
	time	-0.249	0.036	345.611	<0.001 *	S+EX x time	-0.073	0.036	344.087	0.043 *
	time^2	0.013	0.004	326.159	0.005 *	S+EX x time^2	0.010	0.004	326.336	0.028 *
	time^3	<0.001	<0.001	325.116	0.106	S+EX x time^3	<0.001	<0.001	325.289	0.027 *
2. Feelings of Guilt	intercept	1.164	0.622	125.803	0.064	S/EX	0.005	0.073	194.898	0.944
	time	-0.099	0.027	343.249	<0.001 *	S+EX x time	-0.026	0.027	341.507	0.334
	time^2	0.006	0.003	327.526	0.070	S+EX x time^2	0.002	0.003	327.678	0.551
	time^3	<0.001	<0.001	326.602	0.195	S+EX x time^3	<0.001	<0.001	326.795	0.673
7. Work and Activities	intercept	0.533	0.880	121.866	0.546	S/EX	-0.108	0.089	195.226	0.227
	time	-0.269	0.034	330.467	<0.001 *	S+EX x time	-0.053	0.034	328.624	0.120
	time^2	0.020	0.004	316.688	<0.001 *	S+EX x time^2	0.006	0.004	316.853	0.171
	time^3	<0.001	<0.001	315.702	<0.001 *	S+EX x time^3	<0.001	<0.001	315.912	0.190
8. Retardation	intercept	-0.348	0.586	126.323	0.553	S/EX	0.070	0.072	202.526	0.332
	time	-0.001	0.028	347.626	0.985	S+EX x time	-0.094	0.028	345.870	0.001 *
	time^2	-0.006	0.004	332.732	0.105	S+EX x time^2	0.012	0.004	332.888	0.001 *
	time^3	<0.001	<0.001	331.579	0.044 *	S+EX x time^3	<0.001	<0.001	331.781	0.001 *
9. Agitation	intercept	2.978	0.669	129.564	<0.001 *	S/EX	0.032	0.071	206.508	0.651
	time	-0.110	0.028	339.677	<0.001 *	S+EX x time	-0.029	0.028	337.887	0.314
	time^2	0.008	0.004	325.086	0.034 *	S+EX x time^2	0.001	0.004	325.249	0.749
	time^3	<0.001	<0.001	324.019	0.117	S+EX x time^3	<0.001	<0.001	324.223	0.981
10. Anxiety - Psychic	intercept	2.391	0.993	123.756	0.018 *	S/EX	0.063	0.094	202.360	0.502
	time	-0.166	0.037	342.019	<0.001 *	S+EX x time	-0.073	0.037	340.322	0.046 *
	time^2	0.008	0.005	328.841	0.075	S+EX x time^2	0.007	0.005	329.020	0.142
	time^3	<0.001	<0.001	327.702	0.319	S+EX x time^3	<0.001	<0.001	327.917	0.203
11. Anxiety - Somatic	intercept	2.669	0.743	124.506	<0.001 *	S/EX	-0.049	0.081	202.513	0.549
	time	-0.146	0.032	343.753	<0.001 *	S+EX x time	-0.012	0.031	342.046	0.700
	time^2	0.007	0.004	328.722	0.071	S+EX x time^2	-0.001	0.004	328.894	0.791
	time^3	<0.001	<0.001	327.686	0.353	S+EX x time^3	<0.001	<0.001	327.891	0.654

12. Somatic Symp. GI	intercept	-0.305	0.460	124.118	0.509	S/EX	-0.082	0.056	212.388	0.146
	time	-0.053	0.022	347.586	0.019 *	S+EX x time	-0.016	0.022	345.808	0.468
	$time^2$	0.001	0.003	335.005	0.663	S+EX x $time^2$	0.003	0.003	335.148	0.254
	$time^3$	<0.001	<0.001	333.749	0.909	S+EX x $time^3$	<0.001	<0.001	333.952	0.181
13. Somatic Symp. General	intercept	0.270	0.534	125.480	0.614	S/EX	0.081	0.058	205.829	0.164
	time	-0.099	0.023	345.285	<0.001 *	S+EX x time	-0.009	0.023	343.556	0.711
	$time^2$	0.007	0.003	330.824	0.010 *	S+EX x $time^2$	0.001	0.003	330.985	0.644
	$time^3$	<0.001	<0.001	329.700	0.042 *	S+EX x $time^3$	<0.001	<0.001	329.901	0.558
14. Genital Symptoms	intercept	0.106	0.434	119.524	0.807	S/EX	-0.048	0.050	185.435	0.338
	time	-0.050	0.018	346.182	0.005 *	S+EX x time	0.022	0.018	344.405	0.222
	$time^2$	0.004	0.002	332.617	0.070	S+EX x $time^2$	-0.003	0.002	332.769	0.155
	$time^3$	<0.001	<0.001	331.733	0.134	S+EX x $time^3$	<0.001	<0.001	331.935	0.144
15. Hypochondriasis	intercept	0.618	0.964	120.115	0.522	S/EX	-0.131	0.102	168.541	0.199
	time	-0.102	0.033	338.726	0.002 *	S+EX x time	-0.029	0.032	337.119	0.367
	$time^2$	0.006	0.004	320.924	0.134	S+EX x $time^2$	0.004	0.004	321.080	0.270
	$time^3$	<0.001	<0.001	320.427	0.313	S+EX x $time^3$	<0.001	<0.001	320.599	0.257
16. Loss of Weight	intercept	0.249	0.339	122.722	0.464	S/EX	-0.050	0.047	194.249	0.289
	time	-0.096	0.018	345.733	<0.001 *	S+EX x time	0.006	0.018	344.021	0.739
	$time^2$	0.008	0.002	328.414	0.001 *	S+EX x $time^2$	-0.001	0.002	328.568	0.819
	$time^3$	<0.001	<0.001	327.270	0.007 *	S+EX x $time^3$	<0.001	<0.001	327.466	0.826
17. Insight	intercept	-0.035	0.671	126.326	0.959	S/EX	0.032	0.071	202.493	0.654
	time	-0.002	0.028	342.531	0.940	S+EX x time	-0.032	0.027	340.729	0.241
	$time^2$	-0.001	0.003	328.563	0.690	S+EX x $time^2$	0.005	0.003	328.732	0.134
	$time^3$	<0.001	<0.001	327.554	0.677	S+EX x $time^3$	<0.001	<0.001	327.765	0.103

Figure S5. Changes in severity of depressed mood (Item 1)

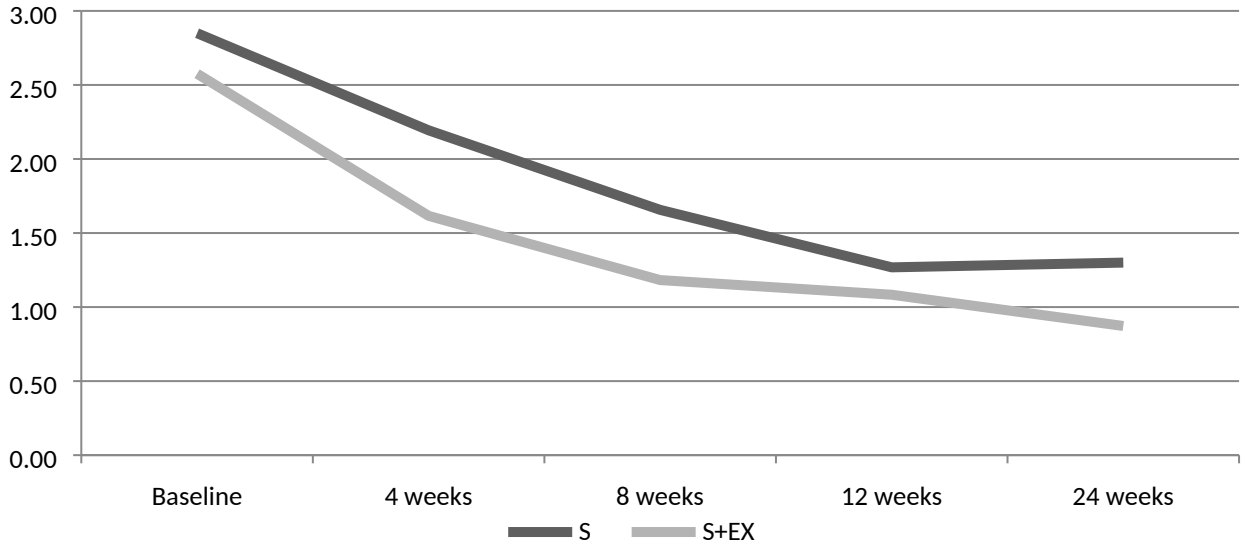


Figure S6. Changes in the severity of psychomotor retardation (Item 8)

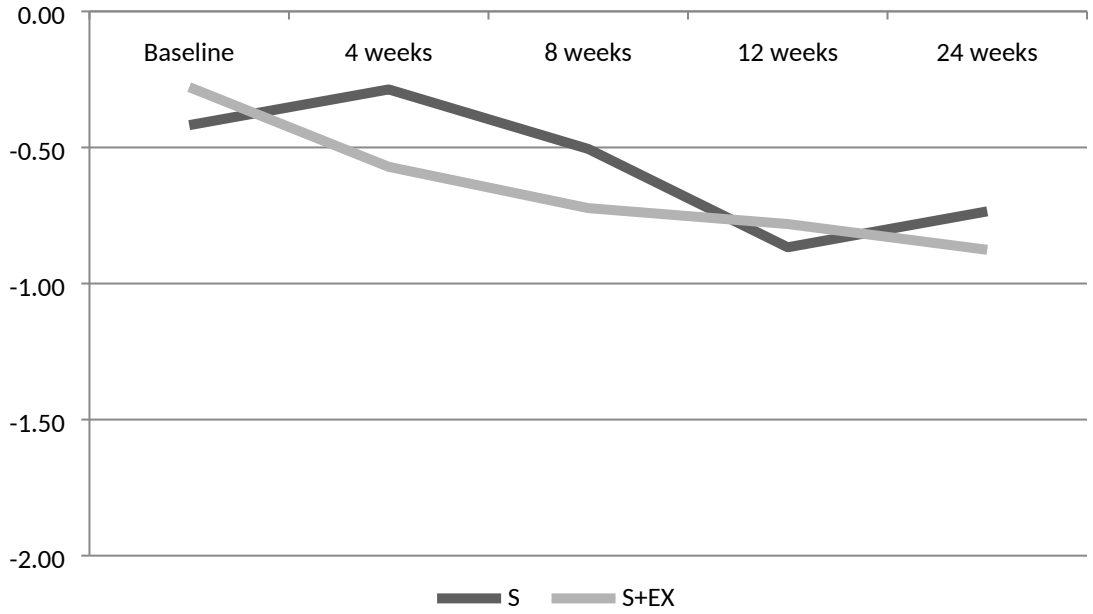
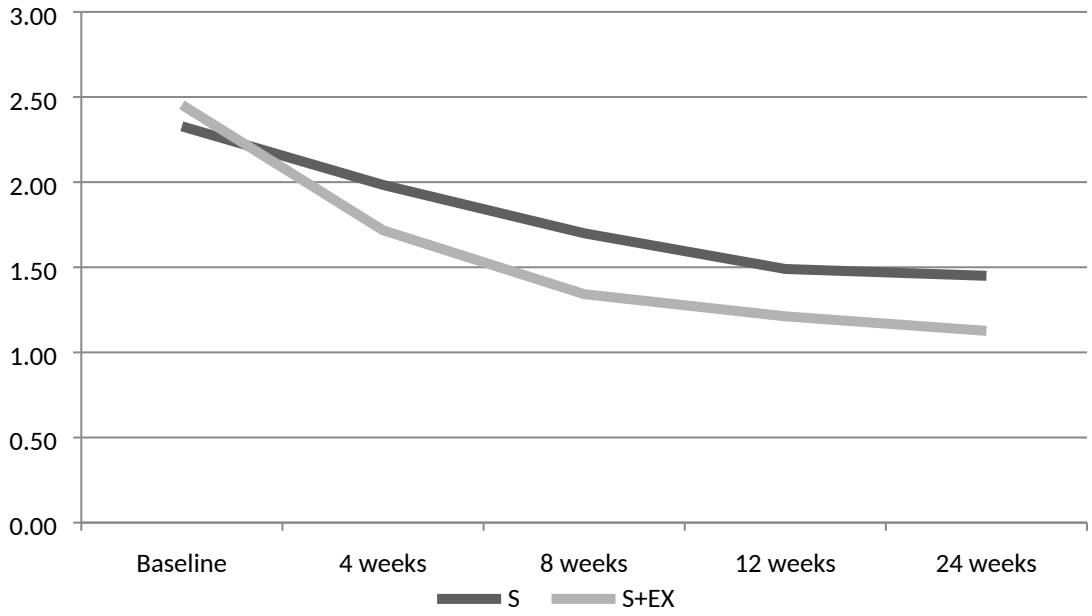


Figure S7. Changes in the severity of psychic anxiety (Item 10)



Appendix. Protocols of physical exercise interventions.

Protocol for the Non-Progressive Exercise intervention (NPE)

Each session of the protocol consisted of:

- 10 min warm-up: walking, strengthening exercises, quiet calisthenics
- 2 repetitions of 10 min each: mat work: stretching, calisthenics, breathing exercises
- 2 repetitions of 5min each: instrumental exercises (first with a ball, then with a stick)
- 2 repetitions of 5min each: balance exercises (e.g. toe walking, heel to toe, single limb stance, staggered stance)
- 10 min cool down: walking, quiet calisthenics

Participants were invited to rest when their heart rate exceeded the threshold of 70% of peak heart rate, or whenever they felt exhausted.

Protocol for the Progressive Aerobic Exercise intervention (PAE)

Each session of the protocol consisted of a 10min warm-up: breathing exercises, slow cycling. This was followed by cycling at an intensity that would maintain the heart rate within the assigned target heart-rate range. Target heart rate was defined by percentage of the peak heart rate (PHR) as measured during the maximum oxygen uptake test. All exercise sessions concluded with 5–10 min of cool-down cycling.

Cycling sessions

- First period (weeks 0–4): exercise bike, cycling at 60–70% of PHR, 30–40min
- Second period (weeks 5–8): treadmill exercise at 70–80% of PHR, 40–50min
- Third period (weeks 9–12): five interval training sessions of 5 min at 85% of PHR or 40 min of continuous treadmill at 70% of PHR
- Fourth period (weeks 13–24): five interval training sessions of 6 min at 85% of PHR, or 40 min of continuous treadmill at 70% of PHR