Bidirectional Prospective Associations Between Physical Activity and Depressive Symptoms. The TRAILS Study

Nikolaos Stavrakakis, M.Sc.*, Peter de Jonge, Ph.D., Johan Ormel, Ph.D., and Albertine J. Oldehinkel, Ph.D.

Interdisciplinary Centre for Psychiatric Epidemiology and Groningen Graduate School, Medical Sciences, Department of Psychiatry, University Medical Centre Groningen, University of Groningen, The Netherlands

Article history: Received March 23, 2011; Accepted September 6, 2011

Keywords: Physical activity; Depressive symptoms; Affective symptoms; Adolescents; Longitudinal studies

ABSTRACT

Purpose: Low levels of physical activity (PA) have been shown to be associated with depression in adults. The few studies that focused on adolescents yielded mixed and inconsistent results. Efforts to examine the direction of this relationship have been inconclusive up to now. The aims of this study were therefore to investigate (1) the direction of the inverse association between PA and depressive symptoms over time, and (2) whether these associations are specific to particular clusters of depressive symptoms in adolescents.

Methods: Depressive symptoms and PA were assessed in a population sample of adolescents (N = 2,230) who were measured at three waves between age 10 and age 17. Depressive symptoms were measured by the Affective Problems scale of the Youth Self-Report and Child Behavior Checklist, whereas PA was operationalized as the amount of time spent on physical exercise. Structural equation modeling was used to examine bidirectional effects of PA and depressive symptoms over time.

Results: We found significant cross-lagged paths from prior PA to later depression as well as from prior depression to later PA (beta values = −0.039 to −0.047). After subdividing depression into affective and somatic symptoms, the affective symptoms were reciprocally related to PA, whereas the paths between somatic symptoms and PA did not reach statistical significance.

Conclusions: An inverse bidirectional association between PA and general depressive symptoms was observed. This association was restricted to affective symptoms.

© 2012 Society for Adolescent Health and Medicine. All rights reserved.
have shown a consistent inverse association between PA and depressive symptoms, results from prospective studies have been inconsistent and will be briefly described later.

Why are there so many mixed results regarding associations between PA and depression? Part of the discrepancies can be attributed to methodological limitations, such as small sample sizes and, for prospective studies, short follow-up periods. Another reason might be that depression is a heterogeneous disorder, consisting of both affective and somatic symptoms. Instruments that are used to assess depressive symptoms vary widely with respect to the symptom profiles they represent, where some instruments do not include somatic symptoms, whereas other instruments contain multiple questions on somatic symptoms. It has been shown that somatic and affective symptoms are differentially associated with cardiac autonomic [14] and HPA axis function [15]. As these two subgroups of symptoms show distinct patterns of association with various physiological measures, they may be differentially related to PA as well.

Despite these discrepancies, several, not necessarily mutually exclusive, hypotheses have been proposed to explain the inverse association between PA and depression:

1. Protection hypothesis (direct causal effect): PA decreases depressive symptoms through biological (elevation of endorphins, serotonin, or endocannabinoids) or social (social contact and self-esteem) mechanisms [16].

2. Inhibition hypothesis (reverse causality): depression negatively influences patterns of PA [17] through symptoms, such as lack of energy, anhedonia, low mood, and social withdrawal [18].

3. Common cause hypothesis: PA and depression share risk factors, for instance genetic or familial factors, such as socioeconomic status (SES), neighborhood, and parental rearing style [19].

In the present study, associations between PA and depressive symptoms were studied in a large prospective population cohort of Dutch adolescents. Studying this association in adolescents is important for at least two reasons. First, evidence regarding the association between PA and depression is even more ambiguous in adolescents than in adults. Although cross-sectional and intervention studies in adolescents have shown moderate to strong associations between PA and depression, respectively [20,21], longitudinal studies have yielded mixed results, varying from strong effects [17,22] to weak [9,23] or even no effects at all [8,24]. Second, a substantial number of adolescents evince subthreshold depressive symptoms [25]. Because adolescent depressive symptoms are thought to predict full-blown depressive disorders later in life [25], researching depressive symptoms in adolescence may help to improve the understanding of the etiology of depressive disorders and help design effective interventions for treating affective disorders.

The first aim of this study is to investigate patterns in the association between PA and depressive symptoms to determine whether changes in PA precede, follow, or co-occur with changes in depressive symptoms in healthy adolescents. A longitudinal design is the only way to illuminate such patterns. The second aim is to examine the association between PA and two subgroups of depressive symptoms, that is, affective symptoms (such as, depressed mood and loss of interest) and somatic symptoms (such as, sleep disturbances and lack of energy). To the best of our knowledge, PA has not been investigated in relation to these symptom groups before.

**Methods and materials**

**Design**

Data for this study were collected as part of the Tracking Adolescents Individual Lives Survey (TRAILS), a prospective cohort of Northern Dutch adolescents. The study was approved by the Dutch Central Committee on Research Involving Human Subjects. A detailed description of its objectives and main design, as well as of the sample selection procedure and nonresponse, can be found in other studies [26].

Data collection was obtained through three assessment waves (T1, T2, T3), which was scheduled from March 2001 to July 2002, September 2003 to December 2004, and September 2005 to August 2008, respectively. Parental written informed consent was obtained after the procedures had been fully explained. Adolescents gave written informed assent at the second and third assessment waves.

**Participants**

Of 2,935 adolescents initially approached, 2,230 (76%; females = 51%, mean age = 11.11, SD = .55) took part in the first assessment wave (T1). The response rate at the second wave (T2) was 96.4% (N = 2,149; 51% females, mean age = 13.65, SD = .53), whereas the response rate at the third measurement wave (T3) was 81.4% (N = 1,816; 53.2% females, mean age = 16.27, SD = .73).

**Instruments**

**Depressive symptoms.** Depressive symptoms were assessed by the Affective Problems scale of the Youth Self-Report (YSR [27]) that was completed by the adolescents at school. Parents filled out the parent version of the YSR, the Child Behavior Checklist (CBCL [28]), at home. The mean scores of the YSR and CBCL scales were used in the analyses. The YSR and CBCL are composed of a list of problems, which are scored on a 3-point scale (0 = never or not at all true, 1 = sometimes true, and 2 = very often or very true). The Affective Problems scale contains 13 items covering depressive symptoms according to the DSM-IV [29], that is, sadness, loss of pleasure, crying, self-harm, suicidal ideation, feelings of worthlessness, guilt, loss of energy, overtiredness, eating problems, and sleeping problems. Because a previous study in the same sample had shown that omission of one sleep item (“I sleep more than most kids”) increased the internal consistency of the scale [15], this item was excluded. Scores on the remaining 12 items were averaged to construct a total depressive symptoms scale with an internal consistency (Cronbach alpha) of .75 at T1, .84 at T2, and .89 at T3. In addition, we constructed a somatic symptoms subscale and an affective symptoms subscale, as described by Bosch et al [15]. The affective symptom subscale included loss of pleasure, crying, self-harm, suicidal ideation, feelings of worthlessness, feelings of guilt, and sadness, whereas the somatic symptoms subscale included lack of appetite, overtiredness, reduced sleep, trouble sleeping, and lack of energy. The internal consistency for the affective symptoms scale was .72 at T1, .82 at T2, and .85 at T3, whereas the consistencies of the somatic symptoms scale were .57 at T1, .69 at T2, and .77 at T3.

**Physical Activity.** At T1, PA was assessed by the question, “How often per week do you perform physical exercise (for example, swimming, playing football, horse-riding)?” The question could
be answered on a 5-point scale (0 = never, 1 = once per week, 2 = two to three times per week, 3 = four to five times per week, and 4 = five to six times per week). At T2 and T3, PA was measured by two questions assessing the time per week the individual engaged in PA during the summer (question 1) and the winter (question 2). “How many days in an average week in the summer/winter do you take part in physical activities?” These questions were rated on an 8-point scale (0 = never to 7 = seven days per week). The questions on summer and winter were averaged, and mean scores were used in the analysis. To achieve a similar metric for the PA measures at all measurement waves, we recorded the T1 data (range, 0–5) into an 8-point scale using the monotonic R7 transform recommended by Little [30].

Confounders

Gender and SES were included as confounders. SES was calculated by averaging five standardized variables (professional occupation and educational attainment for both father and mother, and household income). Three SES groups were created, in which the lowest 25% of scores were categorized as “low SES,” the highest 25% as “high SES,” and the remaining scores were grouped as “intermediate SES.”

Imputation of missing data

A detailed account of attrition rates within the TRAILS study can be found elsewhere [31]. We performed multiple imputations to complete the data set (i.e., N = 2,230 at all three assessment waves) using STATA version 10 (STATA Corp., College Station, TX) and created five different datasets [32], which were subsequently imported into Mplus 5 [33] and were used in the analyses.

Analysis

Cross-sectional and prospective associations between PA and depressive symptoms were investigated by structural equation modeling, using Mplus 5 [33]. As the depression variables (both total depressive symptoms score and scores on the two subscales) were positively skewed and slightly kurtotic, they were log-transformed before analysis. A cross-lagged path model was used to investigate cross-sectional and prospective associations between PA and depressive symptoms. Cross-lagged panel designs take into account the time precedence and control for multivariate dependencies of the antecedent predictor variables. Therefore, besides the cross-lagged paths (interpreted as a link-age of the level of one variable at the first wave with a relative change in another variable in the subsequent assessment wave) between depression and PA and vice versa (Figure 1, path c), autoregressive paths (interpreted as relative stability over time) within depression and PA (Figure 1, path b), and cross-sectional covariances (interpreted as correlations at T1 and as correlated change at T2 and T3) between PA and depression (Figure 1, path a) were also estimated. Correlated change and cross-lagged paths reflect longitudinal relationships.

To test the predictive relationships between PA and depressive symptoms, we developed two cross-lagged models. The first model concerned associations between PA and the total depressive symptoms score, whereas the second model concerned associations between PA and the subscales of somatic and affective symptoms (adjusted for each other). Both models were corrected for gender and SES. As a first step, all cross-lagged and cross-sectional paths were constrained to be equal over time, and the fit indices of the constrained model were compared with those of the unconstrained model. If the fit indices were adequate and not significantly worse than those of the unconstrained model, the more parsimonious (i.e., constrained) model was used. The significance of the cross-lagged paths was established by testing whether removal of the paths led to a deterioration of the model fit. To test whether the two subgroups of depressive symptoms (somatic and affective symptoms) differed with regard to their cross-sectional and cross-lagged association with PA, we compared a model where the paths between somatic symptoms and PA and between affective symptoms and PA, respectively, were constrained to be equal with a model where these paths were allowed to differ.

A good model fit was defined when the comparative fit index (CFI) and the Tucker-Lewis index (TLI) were greater than .95, whereas the root mean square error of approximation (RMSEA) was lower than .05. Ideally, the $\chi^2$ should be nonsignificant ($p > .05$) as well, but larger samples increase the likelihood of obtaining significant $p$ values [34] (Figure 1).

Results

Descriptive statistics

Descriptive statistics of the main variables used in this study are shown in Table 1. Depressive symptoms increased slightly over time, mainly due to a rise in somatic symptoms. PA remained stable over time.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Means (standard deviations) of the main variables used in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1 (mean age, 11.1)</td>
</tr>
<tr>
<td>Total depressive symptoms</td>
<td>.25 (.20)</td>
</tr>
<tr>
<td>Affective symptoms</td>
<td>.21 (.20)</td>
</tr>
<tr>
<td>Somatic symptoms</td>
<td>.32 (.27)</td>
</tr>
<tr>
<td>Physical activity</td>
<td>3.38 (2.09)</td>
</tr>
</tbody>
</table>

N = 2,230 for all measurements since missing data were imputed.

Figure 1. Model showing cross-sectional paths: (a), autoregressive (b), and cross-lagged longitudinal (c) paths. PA = physical activity; T1 = first measurement wave; T2 = second measurement wave; T3 = third measurement wave.
Relationships between depression and PA

Preliminary analyses revealed that cross-lagged paths and correlated changes were constant over time: the model fit parameters of the models where these paths were constrained to be equal across time were adequate, and the \( \chi^2 \) values of the constrained models were not significantly worse than those of the unconstrained model (\( \chi^2 \) difference = 5.555, \( d.f = 3, p > .05 \) for the model involving total depressive symptoms; \( \chi^2 \) difference = 6.365, \( d.f = 6, p > .05 \) for the model involving depression subgroups). Because the constrained models were more parsimonious, we used these models in all subsequent analyses.

Relationships between the total depressive symptoms score and PA are shown in Figure 2 (overall model fit, \( \chi^2 = 13.254, p < .05, d.f = 6 \)). Fit indices showed an excellent fit of model (CFI = .997, TLI = .988, RMSEA = .022), which was significantly better than the fit of the model without cross-lagged paths from PA to depression (\( \chi^2 \) difference = 8.103, \( d.f = 1, p < .05 \)) and the model without cross-lagged paths from depression to PA (\( \chi^2 \) difference = 11.556, \( d.f = 1, p < .05 \)). As expected from these results, the cross-lagged paths were all significant. The autoregressive estimate of depression was higher than that of PA. The cross-sectional association between PA and depression was significant at T1 but not at later assessment waves (Figure 2).

The second model, investigating the relationship between somatic symptoms, affective symptoms, and PA over time, is shown in Figure 3. Indices of approximate fit showed an excellent fit of model variables (\( \chi^2 = 24.170, p < .05, d.f = 13, \) CFI = .998, TLI = .992, RMSEA = .020). Although cross-lagged relationships between PA and somatic symptoms did not reach statistical significance, affective symptoms both predicted and were predicted by PA. This is reflected in the tests comparing the models with and without these cross-lagged paths—although removal of the cross-lagged to and from somatic symptoms did not worsen the model fit significantly (\( \chi^2 \) difference = 2.404, \( d.f = 2, p > .05 \)), removal of the paths to and from affective symptoms led to a significant decrease of the model fit (\( \chi^2 \) difference = 26.755, \( d.f = 2, p < .05 \); Figure 3).

Discussion

The results of this study indicate a bidirectional cross-lagged association between depressive symptoms and PA, in which PA precedes a decrease in depressive symptoms and vice versa. Quite surprisingly, the inverse associations between PA and de-
pressive symptoms concerned affective symptoms, such as depressed mood, loss of pleasure, and low self-worth, in particular. Somatic symptoms, such as sleep disturbances, eating problems, and lack of energy, were not associated with PA. Interestingly, we found that PA remained stable over time, a finding which is inconsistent with some previous studies, which reported a steady decrease of PA over time [35].

Limitations and strengths of this study

This study has a number of notable strengths. First, we used a large, population-based sample of adolescents. The study of adolescents has the following two main advantages: (a) the probability of confounding by prior depressive episodes is low at this age, and (b) the prevalence of somatic conditions that may hamper engagement in PA is still relatively low in adolescents compared with older people. Second, the longitudinal design with three measurements across 6 years made it possible to study changes in PA and depressed mood over a long period. Third, the model fit was excellent, modeling the relationship between PA and depressive symptoms in our data with great accuracy. A final strength is the fact that we subdivided depressive symptoms into somatic and affective symptoms, which, to the best of our knowledge, has not been done before in this context. Depression is a heterogeneous disorder, with different symptoms observed in individual differences. The use of subgroups does justice to this heterogeneity and may help to understand individual differences in the association between PA and depression.

There are also several limitations. First, all the measures used were self-reports, and the measurement of PA reliably relied on a single question, unlike validated PA questionnaires (e.g., IPAQ) that include a series of detailed questions of PA involvement. Self-reports tend to be less reliable than more objective measures of PA, such as \( \text{VO}_{2\text{max}} \), heart rate variability, and calculation of metabolic equivalent of tasks by, for instance, accelerometers. In addition, at the first measurement wave of this study, PA was assessed slightly different than that at the last two waves, in which separate questions for winter and summer, and more response categories were used. Furthermore, we did not measure the nature (aerobic–anaerobic, social–isolated, etc.), intensity, or duration of PA, information which may prove useful for a better understanding of the relationship between PA and depressive symptoms. Concerning the nature of PA, for instance, aerobic exercises have been reported to have an effect on depressive symptoms, whereas anaerobic exercises do not [36]. Similarly, differences in the intensity and the duration of the activity may be related to depressive symptoms [37]. A final limitation is that the reliability of the somatic symptoms subscale was relatively low, which prevents strong associations with other variables.

The associations observed were bidirectional. The more times per week adolescents exercised, the more likely they were to report a reduction in depressive symptoms. The opposite also held true: the more depressive symptoms adolescents exhibited at some point, the more likely they were to report a reduction in PA later on. These findings are in accordance with recent reviews in adults, which showed that regular PA decreases the risk for developing depression [38], and that baseline depression might play an important role in the development of an inactive lifestyle or decreased level of PA [39]. Longitudinal studies have yielded mixed results in adolescents. Although Jerstad et al [9] demonstrated a bidirectional relationship between exercise and depression, Birkeland et al [8] could not. The bidirectional association between PA and depression found in our study provides support for both the protection hypothesis and inhibition hypothesis, which were mentioned earlier in the Introduction section.

We also showed that PA is differentially associated with subgroups of depressive symptoms; the effects are stronger for affective than for somatic depressive symptoms. This lack of association between PA and somatic symptoms is unexpected. Although it seems plausible that adolescents who exercise more, sleep better, eat better have more energy, this was not supported by our data. A reason for this lack of association (between PA and somatic depressive symptoms) may be the aforementioned low reliability of the somatic symptoms scale, which prevents strong relationships. The low reliability implies that there is heterogeneity within the somatic symptom domain. Whether specific somatic symptoms are differentially related to PA was beyond the scope of this article and remains to be investigated in the future.

The inverse associations found between PA and depression were statistically significant but were relatively weak. This is generally in accordance with previous studies which rarely yielded strong effects. This could be due to the measures of PA used, which usually reflect only a small portion of activities and do not take into account the nature of the activities (e.g., voluntary, social, competitive, or isolated), their duration, and their intensity. PA may be beneficial to mental health only after an intensity/duration threshold is reached, which could have weakened the associations. Another likely reason for the weak associations is the fact that both depressed mood and PA are affected by multiple nonshared variables and therefore cannot be expected to explain much variance in each other. Finally, it might be possible that PA and depressive symptoms are not associated in all individuals in the same way; because of genetic or personality differences, some adolescents may benefit from PA, whereas others may not. An elucidation of the underlying mechanisms might provide clues about which individuals may benefit most from PA.

Conclusion

The current prospective study investigated the inverse bidirectional associations between PA and depressive symptoms, in particular, the two subclusters of depressive symptoms (somatic vs. affective), in a population cohort of Dutch adolescents. PA has been shown to be related to depressive symptoms and vice versa but only in relation to affective symptoms. Further research is needed to investigate whether individual symptoms are differentially associated with PA to improve the understanding of this complex relation.

Acknowledgments

This research is part of the TRacking Adolescents’ Individual Lives Survey (TRAILS). Participating centers of TRAILS include various departments of the University Medical Center and University of Groningen, the Erasmus University Medical Center, Rotterdam, the University of Utrecht, the Radboud Medical Center, Nijmegen, and the Parnassia Bavo group, all in the Netherlands. TRAILS has been financially supported by various grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-MW 940 – 38 – 011; ZonMW Brainpower grant 100 – 001 – 004; ZonMw Risk Behavior
and Dependence grants 60–60600–98–018 and 60-60600-97-118; ZonMW Culture and Health grant 261–98–710; Social Sciences Council medium-sized investment grants GB-MaGW 480–01–006 and GB-MaGW 480-07-001; Social Sciences Council project grants GB-MaGW 457–03–018, GB-MaGW 452–04–314, and GB-MaGW 452-06-004; NWO large-sized investment grant 175.010.2003.005; NWO Longitudinal Survey and Panel Funding 481–08–013), the Sophia Foundation for Medical Research (projects 301 and 393), the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP-006), and the participating universities. The authors are grateful to all adolescents, their parents, and teachers who participated in this research and to everyone who worked on this project and made it possible.

References