VU Research Portal

The QLIM study

Braam, K.I.

2016

document version
Publisher's PDF, also known as Version of record

CHAPTER 3

Physical exercise training interventions for children and young adults during and after treatment for childhood cancer

Katja I. Braam, Patrick van der Torre, Tim Takken, Margreet A. Veening,
Eline van Dulmen-den Broeder, Gert-Jan J.L. Kaspers

ABSTRACT

A decreased physical fitness has been reported in patients and survivors of childhood cancer. This is influenced by the negative effects of the disease and the treatment of childhood cancer. Exercise training for adult cancer patients has frequently been reported to improve physical fitness. In recent years, literature on this subject has also become available for children and young adults with cancer, both during and after treatment. This is an update of the original review that was performed in 2011.

This review aims to evaluate the effect of a physical exercise training intervention (at home, at a physical therapy practice, or in-hospital) on physical fitness of children with cancer, in comparison with the physical fitness of children in a care as usual control group. The intervention, with a minimal duration of four weeks, had to be offered within the first five years after diagnosis.

The second aim was to assess the effects of a physical exercise training intervention in this population on fatigue, anxiety, depression, self-efficacy, and health-related quality of life and to assess the adverse effects of the intervention.

For this review the electronic databases of Cochrane Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, CINAHL, PEDro, ongoing trial registries and conference proceedings were searched on 6 September 2011 and updated in 11 November 2014. In addition, a hand search of reference lists was performed in that same period.

The review included randomized controlled trials (RCTs) and clinical controlled trials (CCTs) that compared the effects of physical exercise training with no training, in people who were within the first five years of their diagnosis of childhood cancer.

Two review authors independently identified studies meeting the inclusion criteria, performed the data extraction, and assessed the risk of bias using standardized forms. Study quality was rated by the Grading of Recommendation Assessment, Development and Evaluation (GRADE) criteria.

Apart from the five studies in the original review, this update included one additional RCT. In total 171 participants were included in the analysis, all during treatment for childhood acute lymphoblastic leukemia (ALL).

The duration of the training sessions ranged from 15 to 60 minutes per session. Both the type of intervention and intervention period varied in all the included studies. However, the control group always received usual care.

All studies had methodological limitations, such as small numbers of participants, unclear randomization methods, and single-blind study designs in case of an RCT and all results were of (very) low quality (GRADE).

Cardiorespiratory fitness was evaluated by the 9-minute run-walk test, timed up-and-down stairs test, the timed up-and-go time test and the 20-m shuttle run test. Data of the 9-minute run-walk test and the timed up-and-down stairs test could be pooled. The combined 9-minute run-walk test results showed significant differences between

the intervention and the control group, in favor of the intervention group (standardized mean difference (SMD) 0.69; 95% confidence interval (CI): 0.02 to 1.35). Pooled data from the timed up-and-down stairs test showed no significant differences in cardiorespiratory fitness (SMD -0.54; 95%CI: -1.77 to 0.70). However, there was considerable heterogeneity (I^2 = 84%) between the 2 studies on this outcome. The other 2 single-study outcomes, 20-m shuttle run test and the timed up-and-go test, also showed positive results for cardiorespiratory fitness in favor of the intervention group.

The effect of exercise on bone mineral density (total body) was assessed in one study only, showing a statistically significant positive intervention effect (SMD 1.07; 95% CI: 0.48 to 1.66). The pooled data on body mass index did not show a statistically significant end-score difference between the intervention and control group (SMD 0.59; 95%CI: -0.23 to 1.41).

Flexibility was assessed in 3 studies. Two studies assessed ankle dorsiflexion. One study assessed active ankle dorsiflexion, while the other assessed passive ankle dorsiflexion. No statistically significant difference between the intervention and control group was identified with the active ankle dorsiflexion test; however, in favor of the intervention group, they were found for passive ankle dorsiflexion (SMD 0.69; 95% CI: 0.12 to 1.25). The 3rd study assessed body flexibility using the sit-and-reach distance test, but identified no statistically significant difference between the intervention and control group.

Muscle strength was assessed in 3 studies (knee, ankle, back and leg, and inspiratory muscle strength). Only the back and leg strength combination score showed statistically significant differences on the muscle strength end-score between the intervention and control group (SMD 1.41; 95%CI: 0.71 to 2.11).

Apart from 1 sub-scale of the cancer scale (Worries; P=0.03), none of the health-related quality of life scales showed a significant difference between both study groups on the end-score. For the other outcomes fatigue, level of daily activity, and adverse events (all assessed in one study) no statistically significant differences were found between the intervention and control group.

None of the included studies evaluated activity energy expenditure, time spent on exercise, anxiety and depression, or self-efficacy as an outcome.

The effects of physical exercise training interventions for childhood cancer participants are not yet convincing. Possible reasons are the small numbers of participants and insufficient study designs, but it can also be that this type of intervention is not as effective as in adult cancer patients. However, the first results show some positive effects on physical fitness in the intervention group compared to the control group. Positive intervention effects were seen for body composition, flexibility, cardiorespiratory fitness, muscle strength and health-related quality of life (cancer-related items). As measured by some assessment methods, but not all. However, the quality of the evidence is (very) low and these positive effects were not found for the other assessed

outcomes, such as fatigue, level of daily activity, and adverse events. There is a need for more studies with comparable aims and interventions, using a higher number of participants which also include diagnoses other than ALL.

PLAIN LANGUAGE SUMMARY

Physical exercise training interventions for children and young adults during and after treatment for childhood cancer

Childhood cancer is less common than adult cancer at a rate of 144 to 148 cases per one million children. An intensive treatment, including combined treatment modalities such as surgery, chemotherapy, radiotherapy, or a combination, is often needed for cure. These treatment modalities are frequently accompanied by adverse events, such as nausea, serious infections, organ damage (heart, lung, kidney, liver), decreased bone mineral density, but also decreased muscle strength and physical fitness.

In the past, children were advised to recover in bed, and to take as much rest as possible. Nowadays, it is considered that too much immobility may result in a further decrease of physical fitness and physical functioning. These adverse effects might be prevented or minimized by introducing a physical exercise training intervention during, or shortly after, childhood cancer treatment.

This review includes 5 randomized controlled trials and 1 clinical controlled trial that evaluated the effects of a physical exercise training program in children during cancer treatment. Childhood acute lymphoblastic leukemia (ALL) is the most common type of childhood cancer. For that reason, researchers often focus on this type of cancer since it will provide the largest number of a homogeneous group of patients in the shortest time-span. In total 171 participants with ALL were included in the analysis of this review. The results of the review show that there are some small benefits of physical exercise training on body composition (percentage of fat mass, muscles, and bones), flexibility, cardiorespiratory fitness (endurance capacity), muscle strength and cancer-related health related quality of life, but the evidence is limited. This can be related to an unsuitable intervention for children with cancer, or due to methodological limitations of the included studies. More studies assessing the effects of exercise on body composition, muscle functioning, daily activity, psychological functioning, or a combination of these, are needed in a variety of childhood cancer populations. Furthermore, the current findings do not provide enough evidence to identify an optimal physical exercise training program for children with cancer, neither do they provide information on the characteristics of people who will, or will not, benefit from such a program. These important issues still need to be clarified.

BACKGROUND

Description of the condition

Only a small percentage of the total population suffer from childhood cancer; approximately 144 to 148 cases per million children.[162,163] However the impact of childhood cancer is significant. Many studies report a decreased physical fitness (aerobic capacity and muscle strength), in patients and survivors of acute lymphoblastic leukemia (ALL), which is the most common type of childhood cancer[29,45,63–65,164–168] and also in childhood cancer patients in general[34,53,169–173]. Reduced daily energy expenditure and lower levels of physical activity have been described as the most important cause of this reduced state of physical fitness in childhood cancer patients.[166] In addition, a considerable number of survivors of childhood cancer suffer from motor function disability[17,138], which is mostly related to negative motor signs, such as insufficient muscle activity, or muscle weakness[168,171].

Positive effects of exercise training on physical fitness have been reported in studies with adult cancer patients.[174–177] It is hypothesized that similar results are possible in children with cancer, or survivors of childhood cancer.[64]

Description of the intervention

The intervention under consideration was a physical exercise training program, introduced within the first five years following the diagnosis of childhood cancer. The exercise training should aim to increase physical fitness by aerobic, anaerobic, strength, or mixed fitness training.

How the intervention might work

Cancer and cancer treatment induce lean tissue degeneration and can, therefore, potentially cause abnormalities in the cardiac and skeletal muscle.[178] A decline in protein synthesis and protein degeneration by cancer and its treatment, can reduce muscle mass. This can result in a decreased oxidative enzyme activity and a decreased number of proteins necessary for metabolism.[178] Cancer patients often experience muscle weakness, a decreased functional capacity, decreased flexibility, reduced mobility, and diminished health-related quality of life (HrQoL).[171,178] In addition, a decreased psychosocial functioning and HrQoL as a result of cancer has impact on a person's motivational drive and can result in a poorer self-perception of one's ability to perform physical activity.[166,167]

Physical activity can prevent or diminish the negative effects of a sedentary life-style such as obesity, poor skeletal health, fatigue, and poor mental health, thereby increasing HrQoL of the individual. Increasing physical activity is possible by adopting a less inactive life-style and increasing sports participation. Beneficial effects of physical activity during or shortly after cancer therapy are an increase in muscle mass and plasma volume, improved lung ventilation and lung perfusion, and also an increased cardiac reserve.

This was seen in the study by Dimeo et al (2001); the children with cancer who received cancer treatment with glucocorticoids in combination with resistance exercises, showed less muscle mass loss than the children who did not receive the additional physical exercise training intervention.[179]

Why it is important to do this review

Despite the positive results of exercise interventions on fatigue and physical fitness in adult cancer patients, the evidence for benefits in childhood cancer patients is limited. Studies within the population of childhood cancer patients and survivors are emerging and the first data have recently been published. However, the number of participants in the various publications is small and the variety in type of cancer limited, making it difficult to draw more generalized conclusions. In making healthcare management decisions, participants and clinicians must weigh the benefits and drawbacks of supportive care. Pooled data can help in this decision-making process.

The purpose of this Cochrane review is to summarize the existing literature on the effectiveness of physical exercise training interventions in children with cancer, implemented within the first five years from diagnosis and to provide a best-evidence synthesis or meta-analysis of the reported results. This is an update of the original review that was performed in 2011.[180]

Objectives

Primary objective

To evaluate the effect of a physical exercise training intervention on the physical fitness (i.e. aerobic capacity, muscle strength, or functional performance) of children with cancer within the first five years from their diagnosis (performed either during or after cancer treatment), compared to a control group of childhood cancer patients who did not receive an exercise intervention.

Secondary objectives

To determine whether physical exercise within the first five years of diagnosis has an effect on fatigue, anxiety, depression, self-efficacy, and HrQoL and to determine whether there are any adverse effects of the intervention.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomized controlled trials (RCTs) and controlled clinical trials (CCTs) comparing the effects of physical exercise training within the first five years following the diagnosis of childhood cancer with no training.

A CCT was included in the review when the study included a well-defined and comparable control group. Factors that were taken into account regarding comparability were: being childhood cancer patients or survivors, age, sex, and country of origin.

We included cluster-randomized trials when the intervention and control groups were comparable in each aspect except for the location of cancer treatment and study recruitment.

We included cross-over trials when the study results were available for each separate intervention period. The data of the first randomization period were then used.

Reviews were not included but were checked for relevant references. In addition, we excluded observational studies (including case reports, case-control studies) and surveys from this review.

Types of participants

Study participants were under 19 years of age at diagnosis of any type of childhood cancer. Participants in the physical exercise training program needed to be no more than five years from diagnosis. We only included studies that also included adult cancer participants when the results of the childhood and adult study populations were reported separately.

Types of interventions

Studies that were included compared a physical exercise training intervention for childhood cancer patients or survivors with a control group receiving care as usual. Care as usual is defined as care when needed, but no specific exercise program or alternative intervention prescribed to increase physical fitness, HrQoL, self-perception, or a combination of these, or to decrease adverse events, fatigue, anxiety, depression, or a combination of these in childhood cancer patients.

The physical exercise training interventions that were offered included different types of training or exercise programs. For instance, muscle strength or stretching exercises, aerobic exercises, or sports such as gymnastics, swimming, running, or bicycling.

The exercise training intervention could have been additional care during therapy or could have been offered after the standard cancer therapy in a form of rehabilitation. The goals of this exercise training intervention were preventing motor disabilities and a decline in physical fitness, or treating motor function problems which developed during childhood cancer therapy.

The exercise training intervention could have taken place in any setting or location: at home, at a physical therapy center, in a hospital, or elsewhere. It could either have been a group intervention, or an individual program.

The duration of the exercise training intervention needed to be at least four weeks, in order to be able to report on exercise training effects. The upper limit of the training duration was not fixed for this review. In addition, the duration of physical activities (daily time spent on activities or sports) could differ per protocol.

Types of outcome measures

We included studies evaluating the effect of physical exercise training interventions on physical fitness, HrQoL, fatigue, self-efficacy, anxiety and depression. Furthermore adverse effects of the intervention program were studied.

Primary outcomes

The primary outcome of this review was physical fitness measured by:

1 cardiorespiratory fitness (e.g. peak oxygen uptake (VO_{2peak}), peak work rate (W_{max}), endurance time): aerobic or anaerobic exercise capacity tested by ergometry on a cycle ergometer or treadmill, the Wingate anaerobic test, the steep-ramp-test, maximal anaerobic running/cycling test, the Cooper test, or another valid instrument;

2 muscle endurance/strength: assessed with a hand-held dynamometer, the Biodex, the spring scale, the lateral step-up test, the sit-to-stand test, 10 repetitions maximum, the up-and-down stairs test, the minimum chair height test, the muscle power sprint test, a 10 x 5-m sprint test, the six-minute walk test, the incremental shuttle walking test, or another valid instrument;

3 body composition: using body mass index (BMI), skin-fold measurement, a dual energy x-ray absorptiometry (DXA) scan, waist circumference, or the waist-to-hip-ratio;

4 flexibility: conducted with a goniometer, flexometer or with the sit-and-reach test, V-sit test, shoulder or trunk rotation test, straight leg raise, the passive and active ankle dorsiflexion test, or another valid instrument;

5 activity energy expenditure: for example by using an accelerometer;

6 level of daily activity: assessed by an exercise diary, questionnaire, or by accelerometry;

7 time spent exercising (more than daily activity): assessed by an exercise diary, questionnaire, or by accelerometry

Secondary outcomes

Secondary outcomes of the review were:

1 HrQoL: measured by the Pediatric Quality of Life Inventory (PedsQL), Child Health Questionnaire (CHQ), and DISABKIDS;

2 fatigue: assessed by the PedsQL Multidimensional Fatigue Scale, Childhood Cancer Fatigue Scale (CCFS), or the Fatigue Scale for a child (FS-C), the same scale for adolescents (FS-A), and for parents (FS-P);

3 anxiety and depression: measured by the Childhood Depression Inventory (CDI) and the Center of Epidemiological Studies Depression Scale (CES-D);

4 self-efficacy: assessed using the Confidence Scale, the Self-Efficacy Questionnaire for Children (SEQ-C), or the Children's Self-Efficacy Scale;

5 adverse effects during the study period by collecting information on the occurrence of sport injuries, infections, fractures, heart failure, the recurrence of cancer, and fever.

Search methods for identification of studies

Electronic searches

For this review electronic databases of The Cochrane Register of Controlled Trials (CENTRAL) (*The Cochrane Library*, 11 November 2014, Issue 3), MEDLINE/PubMed (from 1945 to 11 November 2014), EMBASE/Ovid (from 1980 to 11 November 2014), CINAHL (from 1982 to 11 November 2014), and Physiotherapy Evidence Database (PEDro; from 1929 to 11 November 2014) (www.pedro.org.au) were searched.

The search strategies for the different electronic databases (using a combination of controlled vocabulary and text words) are stated in the appendices (see appendix chapter 3).

Searching other resources

We located information about trials not registered in CENTRAL, MEDLINE, EMBASE, CINAHL, and PEDro, either published or unpublished, by searching the reference lists of relevant articles and reviews. We scanned the conference proceedings of the International Society for Pediatric Oncology (SIOP), the American College of Sports Medicine (ACSM), the International Congress on Physical Activity and Public Health (ICPAPH),

and the American Physical Therapy Association (APTA) electronically, or otherwise by hand searching from 2005 till 2014.

A search was performed in the ISRCTN register (www.controlled-trials.com), and the clinical trial database (www.clinicaltrials.gov) for ongoing trials on 11 November 2014. We did not impose language restrictions and will update the searches every two years. The search included "children", "childhood cancer", "cancer", "exercise training therapy", and "outcome" or any related word combination.

Data collection and analysis

Selection of studies

After employing the search strategy described previously, identification of studies meeting the inclusion criteria was undertaken by two review authors (KB, PT) independently. We obtained in full any study that seemed to meet the inclusion criteria on title and abstract, for closer inspection. Reasons for exclusion were noted on a separate form. Discrepancies between review authors were solved by reaching consensus. In one case, a third party arbitrator (TT) was needed: we required another opinion on the study of De Macedo 2010.[181] This discussion resulted in inclusion of that study because the training corresponded with the described criteria of the protocol.

Data extraction and management

Data extraction was performed independently by the two review authors (KB, PT) using standardized forms. For each study we collected information on the study design, participant baseline characteristics, settings, sample size, number of participants in each study arm, type of intervention(s), duration of intervention, randomization and blinding procedure, type of control group, type and duration of cancer treatment and stage of cancer treatment (for example, during or after treatment), and duration of participant follow-up.

The extracted outcome measures included: changes in cardiorespiratory fitness, muscle strength/endurance, body composition, body flexibility, daily energy expenditure per time period (for example, day, week, or month), and changes in the level of daily activity and time spent exercising. In addition, we used a separate form to collect information on psychosocial outcomes such as HrQoL, fatigue, anxiety and depression, and the child's self-efficacy. To collect data regarding any other adverse effect of the intervention, we collected all information reported on adverse events during the intervention period in the included studies. Authors of the studies of which only an abstract was available were contacted for additional study information.

In the process of data extraction consensus was reached on all items.

Assessment of risk of bias in included studies

The two review authors (KB, PT) independently assessed the risk of bias in the included RCTs and CCT. This was done according to the following criteria: random sequence generation (selection bias), allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessor (detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias), and other bias, such as significant baseline imbalance between study groups in pre-score or baseline outcome data. We also looked at differential diagnostic activity to observe differences in study protocol for the intervention group and the control group. For all 'Risk of bias' items of the included studies we used the definitions as described in the *Cochrane Handbook for Systematic Reviews of Interventions*.[182] We included a 'Risk of bias' summary figure. This figure shows whether a study had a high, low, or unclear risk of bias; a green plus symbol corresponds with a low risk of bias, a red minus symbol corresponds with a high risk of bias and the yellow question mark symbol corresponds with lack of information or uncertainty over the potential for bias. Discrepancies between review authors were discussed and solved so consensus was reached. Quality of the outcomes in the different studies was rated by using the Grading of Recommendation Assessment, Development and Evaluation (GRADE) criteria.[183,184] For purposes of systematic reviews, GRADE defines the quality of a body of evidence ('High', 'Moderate', 'Low', or 'Very Low') as the extent to which we can be confident that an estimate of effect or association is close to the quantity of specific interest. The GRADE system entails an assessment of the quality of a body of evidence for each individual outcome.[183] Factors that may decrease the quality of evidence are: 1) study limitations; 2) inconsistency of results; 3) indirectness of evidence; 4) imprecision; and 5) publication bias. Factors that may increase the quality of evidence are: 1) large magnitude of effect; 2) plausible confounding, which would reduce a demonstrated effect; and 3) dose-response gradient.[184] The two review authors performed the quality of evidence grading simultaneously. In case of disagreement they discussed even minor aspects to reach consensus on that matter.

Measures of treatment effect

The main outcome differences between study groups and pooled data are described in the Summary of findings table 3.1. In this table the illustrative comparative risks (with 95% confidence interval (CI)) and differences in standardized mean difference (SMD) are provided. For the Cohen's SMD, data were taken from the post-training/control period measurement. The results of the review also include effect estimates of the intervention per outcome measure. Across the included studies different outcome assessing scales were used. However, in case of BMI we were able to combine data of two studies.

For the interpretation of the Cohen's SMD we used the following criteria[182]:

- less than 0.41 represents a small effect;
- 0.40 to 0.70 represents a moderate effect;
- greater than 0.70 represents a large effect.

Dealing with missing data

Relevant missing data were sought by contacting the primary study author or the corresponding study author. To optimize the strategy for dealing with missing data, we used an intention-to-treat (ITT) analysis when possible. The ITT analysis includes all participants who did not receive the assigned intervention according to the protocol as well as those who were lost to follow-up. Attrition rates, for example dropouts and withdrawals, were investigated to optimize data analyses.

Assessment of heterogeneity

Heterogeneity was assessed both by visual inspection of the forest plots and by a formal statistical test for heterogeneity, that is the I^2 statistic. Significant heterogeneity was defined as $I^2 > 50\%$.[182] In case of heterogeneity, we assessed the following potential sources of clinical heterogeneity: 1) participant characteristics; 2) intervention setting; and 3) stratification methods within studies. When heterogeneity was found, we assessed potential reasons for the differences by examining the study characteristics.

Assessment of reporting biases

In the protocol we had planned to perform a funnel plot, however, due to an insufficient number of studies (fewer than 10) included in this review, we were not able to do so.[182]

Data synthesis

The data of the included studies were entered into Review Manager software (RevMan 2011, version 5.3 Copenhagen: The Nordic Cochrane Centre). The analyses were performed according to the updated *Cochrane Handbook for Systematic Reviews of Interventions*.[182] By using the GRADE criteria, the quality of the included studies was taken into account when interpreting the results for the review. We used the random-effects model throughout the review. When we were unable to perform meta-analysis, we provided all available effect information from the articles.

Subgroup analysis and investigation of heterogeneity

We planned to perform subgroup analyses to evaluate whether the outcome was influenced by differences in the age of the participant, the delivered type of physical exercise training intervention, the duration of the exercise training intervention,

the exercise training intervention location, type of childhood cancer, and cancer treatment.

On three review outcomes a meta-analysis could be performed; i.e. on 9-min run walk test, the timed up-and-down stairs test and BMI. Unfortunately, apart from the intervention and control groups, 9-min run walk test, the timed up-and-down stairs test and BMI data were not available for other subgroup characteristics.[64,65,185,186] Therefore no specific subgroup analysis could be performed.

Sensitivity analysis

For those studies that assessed similar outcomes and of which data could be pooled, we performed sensitivity analyses. We assessed whether the outcome would have been different when a study with high or unclear risk of bias would have been excluded from the analyses. This method aimed to assess whether the findings were robust to the decisions made in the process of obtaining them.

RESULTS

Description of studies

Results of the search

Original review in 2011

For the original version of the review[180] the electronic database searches in CENTRAL, MEDLINE, EMBASE, CINAHL, and PEDro, searches in ongoing trial registries and abstract books from SIOP, ACSM, ICPAPH, and APTA revealed a total of 743 references in 2011. Results of the search from the original review.

After removal of duplicates, the search in 2011 resulted in 710 potentially relevant articles. Initial screening of titles and abstracts excluded a further 700 references that did not meet the criteria for inclusion. The ten remaining references were read in full text. Two out of these ten studies were ongoing trials, four studies did not meet all eligibility criteria and were thus excluded and four studies were included.

Reference list tracking led to two additional articles that potentially could be included. One fulfilled the inclusion criteria and was included in the review. Based on the available in the congress proceeding of the second study, it was not possible to decide if the second study was eligible for inclusion.[187] This study was moved to Characteristics of studies awaiting classification (Table 3.2) see Figure 3.1.

Update in 2014

Running the searches for the update in the aforementioned electronic databases, searching the ongoing trial registries, and searching the abstract books from SIOP, ACSM, ICPAPH, and APTA in November 2014 revealed an additional 704 references.

After removal of duplicates, 643 potentially relevant articles remained. Initial screening of titles and abstracts excluded an additional 620 references that did not meet the inclusion criteria. The remaining 23 references were read in full text. Four of these 23 studies were ongoing trials (see Characteristics of ongoing studies Table 3.3), 15 did not meet all eligibility criteria and three out of the 23 studies were congress proceedings[188-190] see Characteristics of studies awaiting classification Table 3.2). Braam et al (2014) combined two congress proceedings from the same study[190,191] and for that reason were taken together in this review. Only one out of these 23 studies could be included in the update of the original review.

Reference list and trial registry tracking led to four additional articles. Two out of these four studies used a study design that did not match the inclusion criteria of this review. For further information see the excluded study section.[98,192] The other 2 studies were trial protocols which were moved to the ongoing trial section (Table 3.3 Characteristics of ongoing studies) (also see Figure 3.1).

In summary, from the original review and the update together we included a total of six studies in this review.

Included studies

Methods

The six studies included in this review were: Hartman et al. (2009), De Macedo et al. (2010), Marchese et al. (2004), Moyer-Mileur et al. (2009), Tanir and Kuguoglu (2013) and Yeh et al. (2011) (Table 3.4).[63-65,76,181,186] Five of these studies were RCTs, and one study used a quasi-experimental study design, making it a CCT.[76] One study performed a power calculation.[65]

Participants

In total of 171 participants were taken into analysis. All patients were diagnosed with childhood ALL and studied during chemotherapy.[63-65,76,181,186] Of the 171 children, 98 were boys, 70 girls,[63-65,76,181,186] gender was not reported in 3 children who dropped-out. The number of children per study was small. Hartman et al. (2009) included the largest number of children (N = 51) in their study, with 26 children in the usual care group and 25 in the intervention group.[65] The 14 children in the study of De Macedo et al. (2010) were divided in nine children who received care as usual and five who received the intervention.[181] Marchese et al. (2004) included 13 children that performed the exercise intervention and 15 who had care as usual.[63] The 13 children analyzed in the
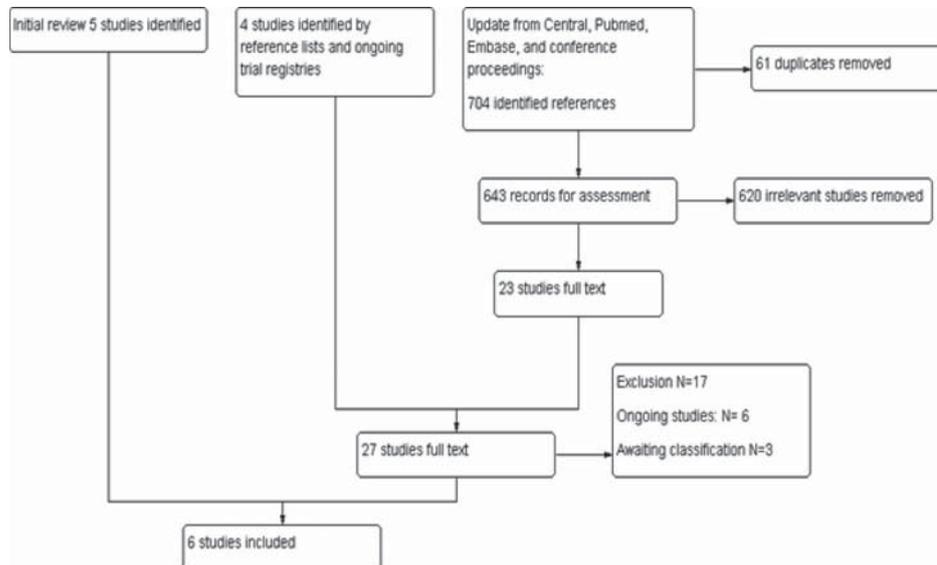
Figure 3.1: Study flow diagram.

study of Moyer-Mileur et al. (2009) were divided in 7 who received care as usual and 6 who received the intervention; 1 child was lost to follow-up.[64] Tanir and Kuguoglu (2013) included 41 children, of which one dropped out, resulting in a group distribution of 19 versus 21 for intervention and control group.[186] Yeh et al. (2011) included 22 children in the analyses, of which 12 children received the intervention training program and 10 received care as usual; 2 children were not taken into analysis because they were lost to follow-up.[76]

Five studies reported their exclusion criteria; in one study these data were missing.[64] Cognitive and/or mental (developmental) impairment was an exclusion criterion in 4 studies.[63,65,76,186] Having difficulties with the national language was described in 1 study.[65] Children with neurological impairment were excluded from participation in 4 studies.[63,76,181,186] Marchese et al. (2004) and Tanir and Kuguoglu (2013) excluded children with a genetic disorder, as well as children who had received cancer-related physiotherapy, or those who had participated in a regular exercise program in a period of six months before start of the study.[63,186] Children with a cardiac, pulmonary, renal or hepatic dysfunction could not participate in the Tanir and Kuguoglu (2013) study,[186] whereas patients with chronic lung disease, neuromuscular disease, or those treated with radiotherapy could not participate in the study of De Macedo et al. (2010)[181].

Intervention

The exercise intervention program of all 6 studies included at least a home-based exercise program with guidance from a therapist of the treating hospital to optimize

physical fitness.[63–65,76,181,186] However, the duration of the entire intervention, the duration of each training session, the timing and the type of the interventions, differed across studies. The duration of the training sessions ranged from 15 minutes up to 60 minutes. The intervention period ranged from 10 weeks[76,181] to two years[65]. Five out of 6 studies introduced the exercise intervention during the maintenance treatment period[63,64,76,181,186] and in one study it started shortly after diagnosis[65]. Five studies determined the effects of an exercise intervention to increase muscle strength of all muscles.[63–65,76,186] The study of De Macedo et al. (2010) investigated the effect of an inspiratory muscle training program.[181] They studied the effects of inspiratory muscle training, which was performed with a threshold device using a load of 30% of the maximal inspiratory pressure (Table 3.4).

Control

The control groups of six studies received care as usual, which implies no additional exercise-related care.[63–65,76,181,186] Noting that Tanir and Kuguoglu (2013) probably made a writing mistake: they reported in the same paper that the control group did and did not receive an exercise intervention.[186] Based on the additional information in the paper we now concluded that the control group did not receive an exercise intervention.

With the exception of those of the study of De Macedo et al. (2010), all study participants of the control groups were measured on the same time points as the intervention group. The control group in the study of De Macedo et al. (2010) performed the study assessments during the initial evaluation and after 10 weeks, whereas the intervention group performed the measurements at the end of each training week.[181]

Outcomes

The studied primary outcomes were: cardiorespiratory fitness, muscle endurance/strength, body composition, flexibility, and level of daily activity. Secondary outcomes of this review that were mentioned in the studies were: HrQoL, fatigue, and adverse events. The other secondary outcomes (anxiety, depression, and self-efficacy) were not addressed.

Because of the different aims and study methods of the six included studies, there was little overlap in used methods and assessed outcomes (Table 3.4). Both the 9-minute run-walk test and the timed up and down stairs test to assess cardiorespiratory fitness were performed in only two studies[63,186] and, changes in BMI, as part of changes in body composition, were studied in another two studies.[64,65]

Table 3.1 Physical exercise training compared with usual care for children and young adults during and after treatment for childhood cancer

Physical exercise training compared with usual care for children and young adults during and after treatment for childhood cancer					
Patient or population: children and young adults during and after treatment for childhood cancer Settings: hospital and non-hospital Intervention: physical exercise training Comparison: usual care					
Outcomes	Illustrative comparative risks* (95% CI)		No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk			
	Usual care group	Exercise group			
Cardiorespiratory outcomes					
9-minute run-walk test wheeled distance counter Follow-up: mean 3-4 months	The mean 9-minute run-walk test in the control group was 3304.5 feet (1007.2 m) and 2676 feet (816 m)	The mean difference between the study groups on 9-minute run-walk test was 681 feet (95% CI: 132 ; 1230) in favor of the intervention group.	68 (2 studies)	⊕⊕⊝⊝ low[1,2]	SMD 0.69 (0.02 to 1.35)
Timed up-and-down stairs test stopwatch Follow-up: mean 3-4 months	The mean timed up-and-down stairs in the control groups was 9.0 sec and 8.6 sec	The mean difference between the study groups on timed up-and-down stairs was -0.94 sec (95% CI: -2.94 ; 1.06) in favor of the intervention group.	68 (2 studies)	⊕⊕⊝⊝ low[1,2,3]	SMD -0.54 (-1.77 to 0.70)
Timed up-and-go Test stopwatch Follow-up: mean 3 months	The mean timed up-and-go test in the control group was 8.3 sec	The mean timed up-and-go test in the intervention group was 6.6 sec (1.3 SD); showing a mean difference of -1.8 (95% CI: -2.7 ; 0.8)	40 (1 study)	⊕⊝⊝⊝ very low[1,2,4]	SMD -1.15 (-1.83 to -0.48)
Body composition outcomes					
Bone mineral density (total body) DXA scan Follow-up: mean 24 months	The control group had a mean standard deviation score (SDS) on (total body) bone mineral density of -1.1 (95% CI: -0.8 to -1.4)	After the 24 months intervention the mean SDS of the intervention group on total body bone mineral density was 0.3 better than in the control group (-0.8 SDS (95%CI: -0.6; -1.2))	51 (1 study)	⊕⊕⊝⊝ low[2]	SMD 1.07 (0.47 to 1.65)
Body mass index Quetelet Index Follow-up: mean 18 months	The mean BMI of the control groups was 1.0 and 0.6.	The intervention group had a mean BMI of 1.2 and 0.7. The mean difference between the study groups on BMI was 0.18 points (95% CI: 0.07 ; 0.30) in favor of the intervention group.	64 (2 studies)	⊕⊝⊝⊝ very low[2,3,4]	SMD 0.59 (-0.23 to 1.41)

Muscle endurance/strength outcomes					
Ankle Dorsiflexion Strength Hand-held dynamometer Follow-up: mean 4 months	The mean ankle dorsiflexion strength in the control groups was **0.22 kg** (normalized for patient weight)	The mean ankle dorsiflexion strength in the intervention groups was **0.25 kg**; showing a mean difference of 0.03 (95% CI: -0.04 ; 0.1)	28 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD 0.29 (-0.46 to 1.04)
Health-related quality of life					
Health-related quality of life PedsQl - General questionnaire (version 3.0) Follow-up: mean 4 months	The mean health-related quality of life in the control groups was **17.5 points**	The mean health-related quality of life in the intervention groups was **15 points**; the mean difference between the study groups was -2.5 (95% CI: -10.1 ; 5.1)	28 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD -0.23 (-0.98 to 0.51)
Fatigue					
General fatigue PedsQl - fatigue questionnaire Follow-up: mean 6 weeks	The mean general fatigue in the control groups was **3.4 points**	The mean general fatigue in the intervention group was 3.3 points. The mean difference between the study groups was -0.15 (95% CI: -3.2 ; 2.9)	22 (1 study)	⊕⊕⊖⊖ low[2,3,5]	SMD -0.04 (-0.88 to 0.8)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in the table. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group. **CI:** Confidence interval; **SMD:** standardized mean difference; **DXA:** dual-energy x-ray absorptiometry

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Footnotes:
1   Quality of the evidence was downgraded 1 level because of high risk of bias in the study of Tanir 2013.
2   Quality of the evidence was downgraded 1 level because of imprecision. All study outcomes were based on small patient numbers. In all studies the total population was less than 400 (a threshold rule-of-thumb value; using the usual α and β, and an effect size of 0.2 SD, representing a small effect) (Schünemann 2009).
3   Quality of the evidence was downgraded because the confidence interval of the SMD crossed no difference (Higgins 2011)
4   Quality of the evidence was downgraded 1 level because for this review outcome the study outcome assessors were not blinded.
5   Quality of the evidence was downgraded 1 level because of high risk of bias in the study of Yeh 2011.

Excluded studies

Twenty-one publications that had been retrieved were subsequently excluded (Table 3.5). There were four studies which included an adult cancer population instead of a pediatric population[193–196] Six studies were excluded based on the used design; one case-control study,[81] one using healthy volunteers as a control population,[98] one used a cross-over randomized trial design without presenting data after the first intervention period (before cross-over),[89] and three were uncontrolled studies[90,192,193]. In another three studies the intervention did not match with the intervention of interest for this review[87,197,198] and one study the aim was to increase motor and process function; this outcome did not correspond with one of the primary or secondary outcomes of this review[199].

Another eight studies assessed the effects of a training intervention with duration of less than four weeks.[80,86,89,200–203] Furthermore, there was duplication of information; within these eight excluded studies (on intervention duration), two studies were described in multiple reports: the first study was reported in Chung (2014) and William Li et al. (2013),[200] the second study results were presented in Speyer et al. (2010) and Speyer et al. (2011[89,203]. The final excluded study was a conference proceeding, presenting data of a pilot study[204]. The full study data were presented separately by Hartman et al. (2009), a study which is included in the review.[65]

The exclusion of two studies was based on two exclusion criteria and therefore mentioned twice in the section abov. [89,193]

Risk of bias in included studies

See Tables 3.4.1 to 3.4.6 and Figure 3.2 for the exact scores per study and the support for the judgments made.

Allocation (selection bias)

The random sequence generation was adequately generated in two out of the six studies (Figure 3.2).[63,65] These two studies used block randomization with sealed envelopes.[63,65] Both De Macedo et al. (2010) and Tanir and Kuguoglu (2013) reported that selection and allocation were random; however, it remained unclear how the randomization procedure was carried out in both studies. A non-randomized design was used in the study of Yeh et al. (2011), leading to a high risk of selection bias. No information on random sequence generation was available for the fifth study.[64] None of the studies described the quality of the envelopes, how the envelopes were sealed, or whether they were coded. Therefore five out of six studies were judged to have an unclear risk of bias for allocation concealment.[63,64,181,186,205] One study did not use a randomization method and therefore had no allocation concealment.[76] In summary,

five studies had an unclear risk of selection bias and due to the absence of a randomization procedure one study had a high risk of selection bias.

Blinding (performance bias and detection bias)

Blinding of participants and personnel (performance bias)
Due to the nature of the interventions blinding was virtually impossible: that is when the participants need to perform an exercise intervention and the children and their parents are well informed about the study purpose, participants cannot be blinded for the study randomization. This could be a potential performance bias in all studies.[182] Therefore, all included studies of this review were thought to have a high risk for performance bias.

Blinding of outcome assessors (detection bias)
It is possible to minimize detection bias with blinding the outcome assessor for the randomization. Two studies used outcome assessors who were blinded for study groups (Figure 3.2).[63,65] In the other four studies the risk was unclear.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Hartman 2009	+	?	-	+	-	+	+
De Macedo 2010	?	?	-	?	?	+	-
Marchese 2004	+	?	-	+	?	+	?
Moyer-Mileur 2009	?	?	-	?	?	+	+
Tanir 2013	?	?	-	?	?	-	-
Yeh 2011	-	-	-	?	+	-	-

Figure 3.2 Risk of bias summary: review authors' judgement about each risk of bias item for each included study

Incomplete outcome data (attrition bias)

All studies reported withdrawals and drop-outs during the intervention period. However, only one study used an ITT analysis to deal with missing data and thus had a low risk of attrition bias.[76]

In the study of Marchese et al. (2004), the authors reported missing data for daily logs of activity and heart monitor.[63] Yet, no information was reported on methods used for data imputation. For two other studies, it also remained unclear whether they used a method for missing data imputation.[64,181] In these 3 studies the risk of attrition bias was therefore unclear.

In the study of Hartman et al. (2009), there was a high risk of attrition bias.[65] The authors used a simple imputation technic to include data for those children who dropped out the study. Yet, they included the data from prior to the elimination. This method is very simple and therefore increases the risk for bias due to incomplete outcome data.

The study of Tanir and Kuguoglu (2013) provided no information about missing data. The authors reported that one child demised in the intervention group. However, data on this child are not presented. In this study the risk of attrition bias therefore was high.

Selective reporting (reporting bias)

In two studies serious selective reporting was detected.[76,186] In the first study results on general quality of life was reported as difference between sexes, instead of difference between the two study groups. In the second study, 'adherence' was mentioned to be an extra, or a secondary outcome. Yet, in the results the authors focused on this item as if it was a primary outcome. In the four other studies the risk of reporting bias was low.

Other potential sources of bias

This review assessed two other biases: baseline outcome data and diagnostic activity. First, the absence of significant differences in baseline outcome data were reported in three studies.[64,65,181] In one study baseline differences on sex, treatment anxiety and goniometer results were reported.[186] In the final two studies it remained unclear whether all baseline test scores were significantly different between the two study groups.[63,76]

Second, the study outcomes were measured at different time points for the intervention and the control group in the study of De Macedo et al. (2010).[181] In the control group the outcomes were assessed during the initial evaluation and after 10 weeks, whereas in the intervention group measurements were performed at the end of each training week. This could have led to differential diagnostic activity. We judged this study to

be of high risk for this other type of bias. The other studies used the same number of measurements, and they were free of 'differential diagnostic activity'.[63–65,76,186]

In summary, two out of the six studies showed unclear the risk of these 'other biases',[63,76] in another two studies the risk was considered high,[181,186] and in the last two studies the 'other bias' risk was low[64,65].

Effects of interventions

Because of the different aims and study methods of the six included studies there was little to no overlap in assessed outcomes. Only three outcomes could be pooled: the timed up and down stairs test, the 9-minute run-walk test (cardiorespiratory fitness) and BMI (body composition).

Cardiorespiratory fitness

In this review cardiorespiratory fitness was defined as: peak oxygen uptake ($VO_{2\ peak}$), peak work rate (W_{max}), or endurance time. In the included studies physical fitness was assessed by the 9-minute run-walk test,[63,186] timed up-and-down stairs test,[63,186] timed up and go test[186] and by the 20-m shuttle run test[64].

The 9-minute run-walk test (SMD 0.69 feet; 95% CI 0.02 to 1.35; P value = 0.04) showed a significant effect in favor for the intervention (N = 32) compared to usual care (N = 36). The analysis showed moderate heterogeneity (I^2 = 44%) for this item between the studies (Figure 3.3.1).

The timed up-and-down stairs test was assessed in 2 studies.[63,186] The test results (SMD -0.54 sec; 95% CI -1.77 to 0.70; P value = 0.40) were not significantly different between the intervention (N = 33) and the control group (N = 36). There was considerable heterogeneity (I^2 = 84%) for this test between the studies (Figure 3.3.2). No ITT analysis could be performed due to missing information on two children who dropped-out the studies.

The timed up-and-go test (SMD -1.15 sec; 95% CI: -1.83 to -0.48) showed a significant positive intervention effect;[186] children of the intervention group were faster in performing the test (Figure 3.3.3).

The 20-m shuttle run test results showed that children who performed home-based exercises during their maintenance chemotherapy for ALL (6 children) were able to reach higher end-scores than those in the control group (7 children) (P value = 0.05) (no RevMan data available). ITT analysis was not performed.[64]

Figure 3.3: Cardiorespiratory fitness outcomes after a physical exercise training intervention for children and adolescents during or after cancer
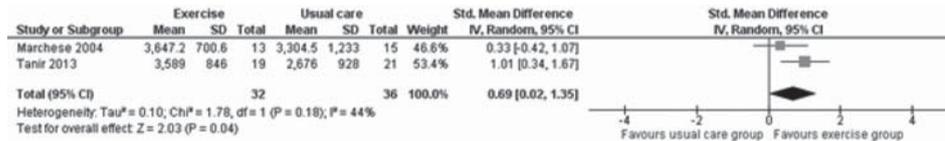
Figure 3.3.1: The 9-minute run-walk test
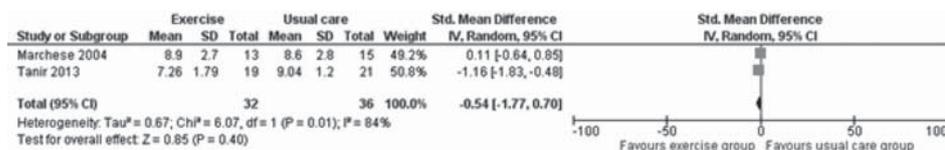
Figure 3.3.2: Timed up-and-down stairs test
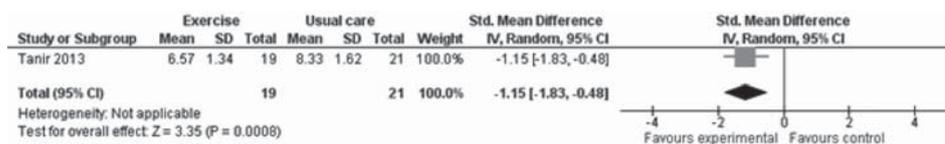
Figure 3.3.3: The timed up-and-go test

Body composition

Bone mineral density (BMD)[65] and BMI[64,65] were assessed as part of the outcome body composition.

The study of Hartman et al. (2009) used a DXA scan to determine BMD (lumbar spine and a total body) changes in childhood ALL participants.[65] The assessments were performed at diagnosis, during chemotherapy, and 1 year after the end of treatment. Analysis showed a significant SMD 1.07 for total body BMD (95% CI 0.48 to 1.66; P value < 0.001) after the intervention of 24 months (Figure 3.4.1). These results revealed a large and significant positive intervention effect on the total body BMD for the intervention group (N = 25) compared to the control group (N = 26). This analysis was performed according to the principles of the ITT analysis.

Differences in BMI between the intervention group and the control group were studied in two trials[64,65] (Figure 3.4.2). In the study of Moyer-Mileur et al. (2009) no intervention effect on BMI was found (SMD 0.02; 95% CI -1.07 to 1.11).[64] This study compared six children who received a combined nutrition and exercise program, with seven children who received usual care. Data of a child that dropped out were not reported and we, therefore, could not perform an ITT analysis on BMI. The study of Hartman et al. (2009) did show a statistically significant difference on BMI in favor of the exercise group (N = 25) compared to the control group (n = 26) (SMD 0.90; 95% CI 0.32 to 1.48).[65] These BMI analyses were performed according to ITT analysis principles.[65] Pooled data analysis for BMI showed a non-significant intervention effect: SMD 0.59 on the Quete-

let Index (95% CI -0.23 to 1.41; P value = 0.16) in favor of the intervention group. In addition, analysis also showed no substantial heterogeneity (I^2 = 48%) for this item between the studies (Figure 3.4.2).

Figure 3.4: Body composition outcomes after a physical exercise training intervention for children and adolescents during or after cancer

Study or Subgroup	Exercise Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
Hartman 2009	-0.8591	0.2778	25	-1.1412	0.2415	26	100.0%	1.07 [0.48, 1.66]	
Total (95% CI)			25			26	100.0%	1.07 [0.48, 1.66]	

Heterogeneity: Not applicable
Test for overall effect: Z = 3.55 (P = 0.0004)

Favours usual care group   Favours exercise group

Figure 3.4.1: Bone minderal density

Study or Subgroup	Exercise Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
Hartman 2009	1.2023	0.2214	25	1.0041	0.212	26	64.4%	0.90 [0.32, 1.48]	
Moyer-Mileur 2009	0.65	0.34	6	0.64	0.445	7	35.6%	0.02 [-1.07, 1.11]	
Total (95% CI)			31			33	100.0%	0.59 [-0.23, 1.41]	

Heterogeneity: Tau² = 0.19; Chi² = 1.94, df = 1 (P = 0.16); I² = 48%
Test for overall effect: Z = 1.40 (P = 0.16)

Favours usual care group   Favours exercise group

Figure 3.4.2: Body mass index

Flexibility

In two studies the ankle dorsiflexion range of motion was measured. However, in one study this was done in a passive way[65] and in the other by active contraction of the ankle[63]. Therefore data could not be pooled.

According to the ITT analysis shown in Figure 3.5, the passive ankle dorsiflexion showed a moderate significant positive effect for the 25 children in the intervention group compared to the 26 children in the control group (SMD 0.69; 95% CI: 0.12 to 1.25; P value = 0.02).[65] Analysis of the ankle dorsiflexion range of motion, measured in active contraction, showed a non-significant moderate effect in the intervention group (13 children) compared to the control group (15 children) (SMD 0.46; 95% CI -0.29 to 1.22; P value = 0.23) (Figure 3.5).[63] Because Marchese et al. (2004) only provided the data of the children who completed all measurements, no ITT analysis was performed.[63]

Study or Subgroup	Exercise Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
3.1.1 Active ankle dorsiflexion									
Marchese 2004	12.5	6.3	13	9.8	5.1	15	100.0%	0.46 [-0.29, 1.22]	
Subtotal (95% CI)			13			15	100.0%	0.46 [-0.29, 1.22]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.20 (P = 0.23)									
3.1.2 Passive ankle dorsiflexion									
Hartman 2009	4.8199	1.2496	25	3.9631	1.2134	26	100.0%	0.69 [0.12, 1.25]	
Subtotal (95% CI)			25			26	100.0%	0.69 [0.12, 1.25]	

Heterogeneity: Not applicable
Test for overall effect: Z = 2.37 (P = 0.02)

Favours usual care group   Favours exercise group

Test for subgroup differences: Chi² = 0.22, df = 1 (P = 0.64), I² = 0%

Figure 3.5: Flexibility outcomes after a physical exercise training intervention for children and adolescents during or after cancer

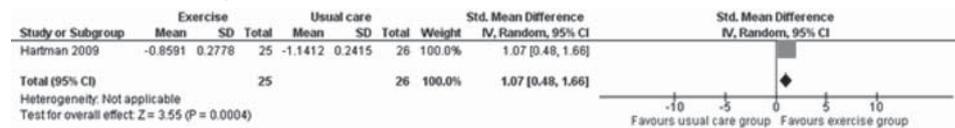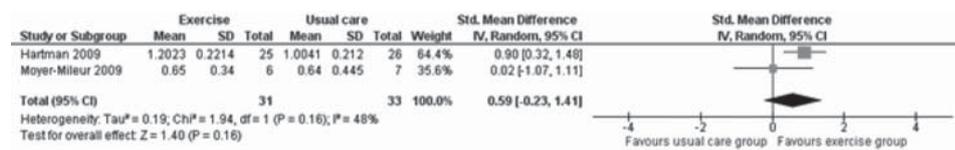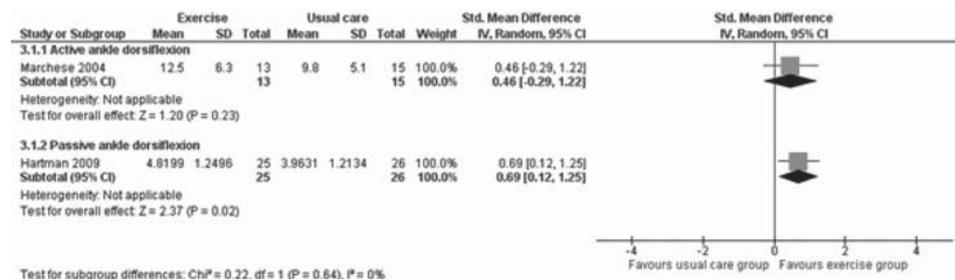The study of Moyer-Mileur et al. (2009) assessed body flexibility with the sit-and-reach distance test.[64] In this study there was no difference in the test results between the 6 children of the intervention and 7 children of the control group. P-values and ITT analysis were not stated in the text or provided by the authors.

In the study of Tanir and Kuguoglu (2013) the goniometer was used to assess the range of motion.[186] Unfortunately, the baseline scores were significantly different. Possible affecting the outcomes. This study showed statistically significant differences between the study groups at the end-measurement with higher scores in the control group, but no significant increase of the goniometer results over time within the study groups. The authors, however, did not report the assessment position of the goniometer on the body. It therefore is not clear whether a decrease in the goniometry results over time is a positive or a negative study result.

Muscle endurance/strength

Marchese et al. (2004) assessed the knee and ankle strength changes and Tanir and Kuguoglu (2013) assessed back and leg strength changes by hand-held dynamometry.[63,186] In both studies the authors found a significant effect in favor of the intervention group. Analysis showed that differences between the end scores of the intervention group and the control group were not significantly different for both knee and ankle strength (Figure 3.6.1 and 3.6.2), but were significant for back and leg strength SMD 1.41 (95% CI 0.71 to 2.11; P value < 0.01) (Figure 3.6.3) Tanir and Kuguoglu (2013). The SMD of the knee strength was 0.25 (95% CI -0.49 to 1.00; P value = 0.51) and the increase of ankle strength was 0.29 (95% CI -0.46 to 1.04; P value = 0.44).[63]

The study of Moyer-Mileur et al. (2009) reported differences in the number of completed push-ups (with knees on the ground) and used a peripheral quantitative computed tomography of the tibia to determine the muscle mass of the participants.[64] According to the original study data, there was no significant change in the maximum number of push-ups or muscle mass, within or between the intervention (6 children) and control group (7 children). The report of this study did not include the data of these results; therefore RevMan analysis could not be performed.

Respiratory muscle strength in the study of De Macedo et al. (2010) was assessed by measuring the maximal inspiratory pressure and maximal expiratory pressure with a digital manometer and a nozzle to dissipate additional pressure caused by the facial muscles and the oropharynx.[181] In the intervention group (5 children) the authors found a significant improvement over time compared to the control group (9 children). Yet, the end score differences were not significant different between the study groups; SMD for inspiratory breathing muscle strength was 0.33 (95% CI -0.77 to 1.43; P value = 0.56), for expiratory breathing muscle strength the SMD was 0.00 (95% CI -1.09 to 1.09; P value = 1.00) (Figure 3.6.4 and 3.6.5).

Figure 3.6: Muscle endurance and strength after a physical exercise training intervention for children and adolescents during or after cancer
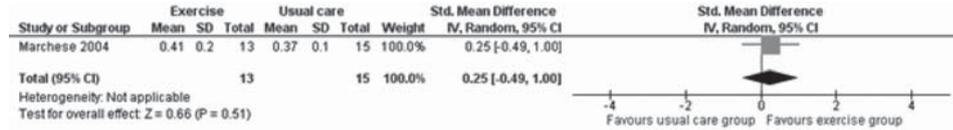
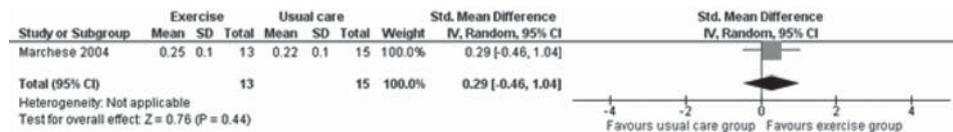| Study or Subgroup	Exercise			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
Marchese 2004	0.41	0.2	13	0.37	0.1	15	100.0%	0.25 [-0.49, 1.00]	
Total (95% CI)			13			15	100.0%	0.25 [-0.49, 1.00]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.66 (P = 0.51)									Favours usual care group Favours exercise group

Figure 3.6.1: Knee strength

Study or Subgroup	Exercise			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
Marchese 2004	0.25	0.1	13	0.22	0.1	15	100.0%	0.29 [-0.46, 1.04]	
Total (95% CI)			13			15	100.0%	0.29 [-0.46, 1.04]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.76 (P = 0.44)									Favours usual care group Favours exercise group

Figure 3.6.2: Ankle strength

Study or Subgroup	Exercise			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
Tanir 2013	75.52	30.22	19	40.71	17.19	21	100.0%	1.41 [0.71, 2.11]	
Total (95% CI)			19			21	100.0%	1.41 [0.71, 2.11]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 3.94 (P < 0.0001)									Favours control Favours experimental
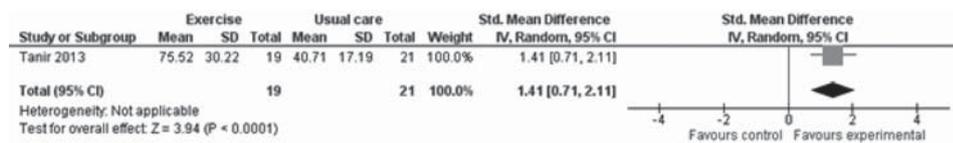
Figure 3.6.3: Back and leg strength

Study or Subgroup	Exercise			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
Macedo 2010	66	21.6	5	59.8	15.4	9	100.0%	0.33 [-0.77, 1.43]	
Total (95% CI)			5			9	100.0%	0.33 [-0.77, 1.43]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.58 (P = 0.56)									Favours usual care group Favours exercise group
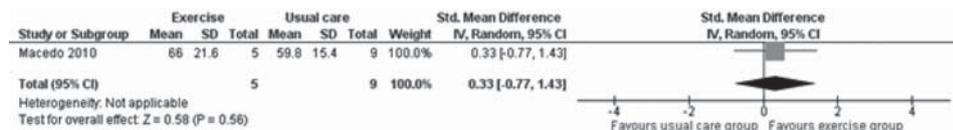
Figure 3.6.4: Inspiratory breathing muscle strength

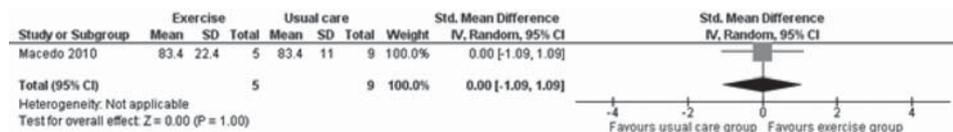| Study or Subgroup	Exercise			Usual care			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
Macedo 2010	83.4	22.4	5	83.4	11	9	100.0%	0.00 [-1.09, 1.09]	
Total (95% CI)			5			9	100.0%	0.00 [-1.09, 1.09]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.00 (P = 1.00)									Favours usual care group Favours exercise group

Figure 3.6.5: Expiratory breathing muscle strength

Due to invalid methods used for missing data imputation, an ITT analysis could not be performed for these outcomes.

Activity energy expenditure
No information was available for activity energy expenditure as it was not assessed in the included studies.

Level of daily activity
Daily physical activity of the participants was assessed in one study.[64] This study used both the pedometer steps-per-day and an activity questionnaire to examine physical activity behavior. This study showed that the increase in 'reported activity in minutes per day' over time was approximately the same for the six children in the

intervention group. In the control group three out of seven children increased in their reported activity in minutes per day. According to the original analyses the reported activities at baseline and at six months were not statistically significantly different between the intervention group and the control group.[64] At 12 months from baseline a higher number of steps was recorded in the intervention group compared with the controls, but this difference was of borderline statistical significance (P value = 0.06) (no RevMan data available).[64] This analysis was not performed according to the ITT procedure.

Time spent exercising (more than daily activity)
No information was available for activity energy expenditure as it was not assessed in the included studies.

Health-related quality of life
General HrQoL was assessed by the PedsQL generic core scale (version 3.0 and 4.0) by the study of Marchese et al. (2004) and in the study of Tanir and Kuguoglu (2013) respectively.[63,186]

Marchese et al. (2004) did not find a significant effect on quality of life effect by the physical exercise training intervention.[63] Overall, the SMD for PedsQL Generic was -0.23 (95% CI -0.98 to 0.51; P value = 0.54) (Figure 3.7.1). The study of Tanir and Kuguoglu (2013) did not report the intervention effects on general HrQoL.[186] For general HrQoL, the authors only provided results for males and females separately at the two time-points of the study.[186] Therefore, pooling of the HrQoL results was not possible.

Cancer-related HrQoL was assessed by the PedsQL cancer module 3.0 in both the study of Marchese et al. (2004) and Tanir and Kuguoglu (2013).

Tanir and Kuguoglu (2013) found an increase on cancer-related HrQoL in both groups by patient-report on: pain and hurt, nausea, and procedural anxiety scales; without significant differences on the end-scores.[186] The only significant different end score between the intervention (19 children) and control group (21 children) was found for the sub-scale assessing worries (in favor of the intervention group).[186]

The SMD on the PedsQL Cancer Module was 0.16 (95% CI -0.58 to 0.91; P value = 0.66) (Figure 3.7.2.). In addition to the patient-reported data, parent reports also showed no intervention effect: the SMD on the parent general PedsQL questionnaire was 0.38 points (95% CI -0.37 to 1.13; P value = 0.32 (Figure 3.7.3) and the SMD on the parent cancer PedsQL was 0.04 points (95% CI -0.70 to 0.79; P value = 0.91) (Figure 3.7.4). Due to missing data an ITT analysis could not be conducted.

Figure 3.7: Health-related quality of life outcomes after a physical exercise training intervention for children and adolescents during or after cancer
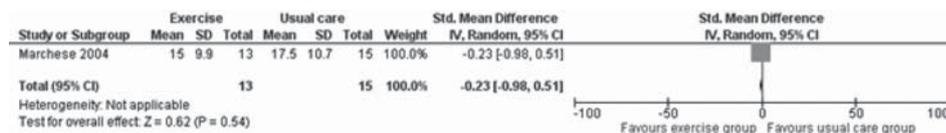
Study or Subgroup	Exercise Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Marchese 2004	15	9.9	13	17.5	10.7	15	100.0%	-0.23 [-0.98, 0.51]
Total (95% CI)			13			15	100.0%	-0.23 [-0.98, 0.51]

Heterogeneity: Not applicable
Test for overall effect: Z = 0.62 (P = 0.54)

Favours exercise group   Favours usual care group

Figure 3.7.1: PedsQL Generic - child report

Study or Subgroup	Exercise Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Marchese 2004	16.4	12.8	13	14.53	9.2	15	100.0%	0.16 [-0.58, 0.91]
Total (95% CI)			13			15	100.0%	0.16 [-0.58, 0.91]

Heterogeneity: Not applicable
Test for overall effect: Z = 0.43 (P = 0.66)

Favours exercise group   Favours usual care group

Figure 3.7.2: PedsQL Cancer Module - child report

Study or Subgroup	Exercise Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Marchese 2004	20.1	11.5	13	16.3	7.9	15	100.0%	0.38 [-0.37, 1.13]
Total (95% CI)			13			15	100.0%	0.38 [-0.37, 1.13]

Heterogeneity: Not applicable
Test for overall effect: Z = 0.99 (P = 0.32)

Favours exercise group   Favours usual care group

Figure 3.7.3: PedsQL Generic - parent report

Study or Subgroup	Exercise Mean	SD	Total	Usual care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Marchese 2004	21.5	14	13	20.9	13.1	15	100.0%	0.04 [-0.70, 0.79]
Total (95% CI)			13			15	100.0%	0.04 [-0.70, 0.79]

Heterogeneity: Not applicable
Test for overall effect: Z = 0.11 (P = 0.91)

Favours exercise group   Favours usual care group

Figure 3.7.4: PedsQL Cancer Module - parent report

Fatigue

Yeh et al. (2011) measured the effect of a physical exercise training intervention on fatigue.[76] This study used the PedsQL multidimensional fatigue scale. They compared changes on fatigue between the intervention group (12 children) and the control group (10 children) over eight time points within ten weeks. There were no significant differences between the intervention and control group on sub-scale general fatigue (SMD -0.04; 95% CI -0.88 to 0.80; P value = 0.92) (Figure 3.8.1). More specifically, no intervention effect was seen for sleep and rest (SMD -0.01; 95% CI -0.85 to 0.83; P value = 0.98) (Figure 3.8.2), nor for cognitive fatigue (SMD 0.07; 95% CI -0.77 to 0.91; P value = 0.86) (Figure 3.8.3). Apart from a per-protocol analysis, the study of Yeh et al. (2011) included an ITT analysis. The ITT analysis revealed no significant intervention effects on fatigue.

Figure 3.8: Fatigue outcomes after a physical exercise training intervention for children and adolescents during or after cancer
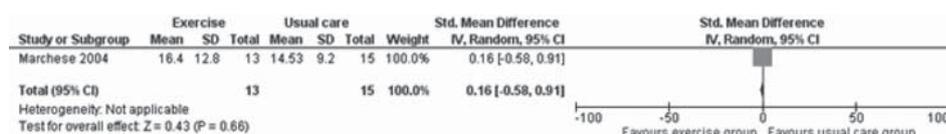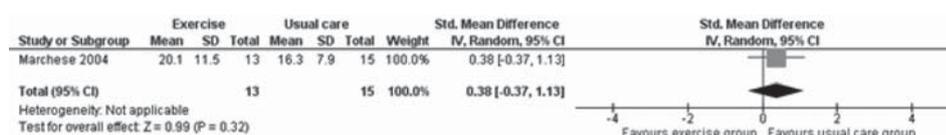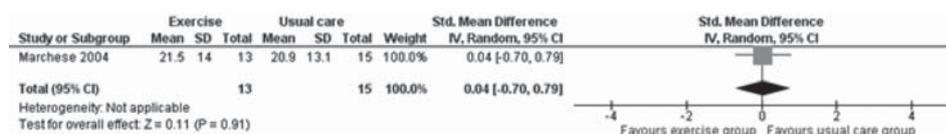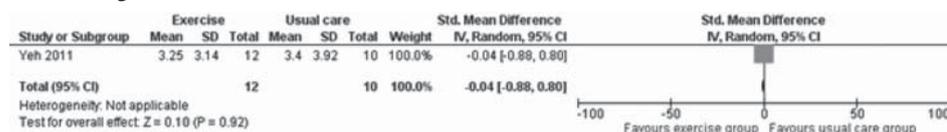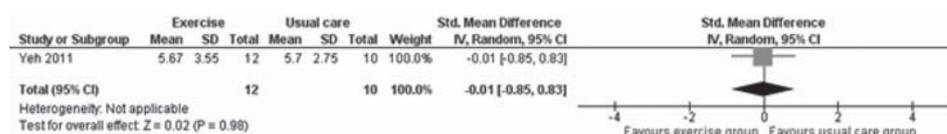

Figure 3.8.1: PedsQL General Fatigue
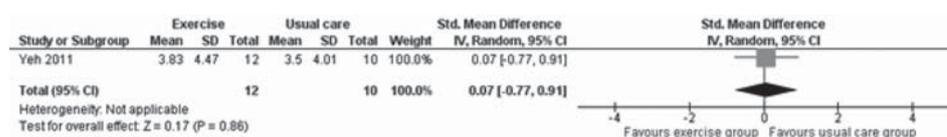

Figure 3.8.2: PedsQL Sleep and Rest

Figure 3.8.3: PedsQL Cognitive Fatigue

Anxiety and depression

No information was available for anxiety and depression as these items were not assessed in the studies included.

Self-efficacy

No information was available for self-efficacy as this item was not assessed in the studies included.

Adverse events (due to, or not clearly related to, the intervention)

The study of Marchese et al. (2004) reported that none of the children experienced any negative effect from the exercises or experienced complications attributed to the physical program. The other studies did not report on this item.[64,65,76,181,186]

Sensitivity analysis

Sensitivity analyses were performed for those outcomes for which pooling was possible (i.e. 9-minute walk-run test, timed up and down stairs test and BMI).[63–65,186] We assessed whether the outcome would have been different when a study with high or unclear risk would have been excluded in the review analyses.

Both the 9-minute-walk-run test and the timed-up-and-down stairs test were performed by Marchese et al. (2004) and Tanir and Kuguoglu (2013).[63,186] In these studies there were three bias items: Random sequence generation (selection bias), blinding

of outcome assessment (detection bias), and selective reporting (reporting bias), in which Tanir and Kuguoglu (2013) showed high or unclear bias compared to low bias in the study of Marchese et al. (2004). For these three items sensitivity analyses were performed. For all other risk of bias items the two studies scored the same (i.e. high or unclear risk) or performed a combination of high and unclear risk.

The outcome of the sensitivity analysis for the 9-minute-walk-run test of Marchese et al. (2004) without the data of Tanir and Kuguoglu (2013) showed a SMD of 0.33 (95% CI: -0.42 to 1.07) whereas the results including Tanir and Kuguoglu (2013) showed a significant intervention effect with a SMD of 0.69 (95% CI: 0.02 to 1.35). This analysis showed the analyses were consistent among the trials.

Additionally, the same was done for the timed-up-and-down-stairs test. When data of the study by Marchese et al. (2004) was assessed without the data of Tanir and Kuguoglu (2013), data showed a non-significant SMD of 0.11 (95% CI: -0.64 to 0.85). Comparable results were found when the data of both studies were included (SMD -0.54; 95% CI: -1.77 to 0.70). The results of the trials were therefore again consistent among the trials.

BMI was assessed in the studies of Hartman et al. (2009) and Moyer-Mileur et al. (2009).[64,65] In these studies there were two bias items: random sequence generation (selection bias) and blinding of outcome assessors (detection bias) on which Moyer-Mileur et al. (2009) showed unclear bias compared to the study of Hartman et al. (2009). For these items sensitivity analyses were possible. For all other risk of bias items the two studies scored the same (i.e. low, high, or unclear risk) or performed a combination of high and unclear risk.

The outcome of the sensitivity analysis showed the BMI data of Hartman et al. (2009) without Moyer-Mileur et al. (2009) (SMD 0.90; 95% CI 0.32 to 1.48). The results of the pooled data revealed a SMD of 0.59 (95% CI -0.23 to 1.41). The results of the sensitivity analyses thus were consistent among the trials and did not differ from the overall analyses.

DISCUSSION

Summary of main results

Several studies have investigated the effects of physical exercise training interventions on physical fitness in adult cancer patients, showing different benefits. Limited studies investigated the effects of such an intervention in a childhood cancer population. Especially high quality studies with an RCT or CCT design are still lacking in this field of research.

This is an update of the original review that was performed in 2011.[180] This updated review include six original studies. All studies investigated the effects of a physical exercise training intervention, with a duration of at least four weeks, in children with cancer. They all aimed to improve physical functioning or psychosocial well-being, and had enrolled children with ALL. However, the studies had two important limitations. Firstly, the total number of participants included in the six studies was limited, and secondly, the exercise programs were not always appropriately designed to meet the study goals.

Cardiorespiratory fitness was studied by the use of the 9-minute run-walk test, the timed up-and-down stairs test, the timed up and go time test and the 20-m shuttle run test. All but the timed up-and-down stairs test showed significant positive intervention effects (P value < 0.05).

Bone mineral density was found to increase significantly higher after a physical exercise training intervention when compared with the study control group. BMI was assessed in two studies. One study found a significant intervention effect on BMI. However, these results were not found when the data were pooled with the second study.

Flexibility was assessed in four studies and each study used different test methods. No (statistically significant) differences between the study groups were identified in three studies, whereas in the fourth study a statistically significant difference in favor of the exercise group was found.

Muscle strength showed a significant intervention effect in one study when assessing back and leg strength. The other three studies assessing muscle strength could not report statistically significant intervention effect on muscle strength. Neither on knee or ankle muscle strength, which were assessed in two studies, nor on lung muscle strength (maximal inspiratory and expiratory pressure) which was the primary outcome of the fourth study.

HrQoL assessed by the cancer module of the PedsQL showed some positive effects in the intervention group in comparison to the control group in one study.[186] No statistically significant differences between the study groups were found for the level of daily activity and fatigue. In addition, only one study reported no complications attributed to the physical exercise intervention program, whereas the other studies did not address this item.

None of the six included studies evaluated the outcomes of activity energy expenditure, time spent exercising, anxiety and depression, or self-efficacy.

It should be noted that the exercise interventions were not the same and the quality and quantity of the evidence was limited.

For future research it is advised to assess the effects of one type of exercise intervention in a larger group of childhood cancer patients, preferably in children with ALL as well as other childhood cancer diagnoses. This can be done in well-designed studies with large sample sizes.

Overall completeness and applicability of evidence

This review provides evidence for modest positive effects of physical exercise training interventions for children with cancer. These modest effects could be due to small sample sizes, various types of interventions provided, and different outcome measures that were used in the six studies. As a result, only data for 9-minute walk-run test, the timed up and down stairs test and BMI could be pooled; therefore, the results of the analysis were instable and weak. However, the meta-analysis and sensitivity analysis on these three outcomes showed consistent results. Furthermore, the patient population was unintentionally homogeneous since; all included children had ALL. The results of this review, therefore, are not applicable for other types of childhood cancer.

The RevMan analyses results of this review are very different to the analysis performed by the authors of some of the studies, which led to different conclusions. For De Macedo et al. (2010), Hartman et al. (2009), and Marchese et al. (2004), the differences were due to different methods of analysis.[63,65,181] In this review we assessed the final outcome differences between the study groups (Figure 3.6.1, 3.6.4 and 3.6.5) and did not asses changes over time.

The included studies all had supervised interventions with a duration and intensity in which it was possible to have a physiological response.[63–65,76,181,186] From literature it is known that supervised exercise interventions in children are more effective compared to non-supervised programs.[206] It is also known that a well-designed exercise program consists of four parameters: mode (type of exercise), intensity, frequency, and duration.[207,208] It would be advisable for new studies to first determine if the planned program includes all elements of these parameters. This will increase the quality of the trials and also increase the comparability.

Appropriate statistical methods are important. The use of incorrect statistical methods can diminish the likelihood of demonstrating the real effects, also in high-quality interventions. In this review only one of the included studies used a power calculation.[65] In the included studies the authors used a Chi^2 test or the Mann-Whitney U test,[64,65,186] the Kruskal-Wallis,[64,186] and the paired sample T-test[181,186] to assess baseline (pre-score) differences between the study groups. The baseline scores were reported as group average,[63,65,76,181,186] but also per study participant[64]. These baseline differences might have had a large impact on the results and conclusions of this review. It would have been preferable for all authors to have corrected for baseline differences in their analyses. However this was not done. To increase the quality of evidence of this review we hoped to be able to pool all raw data (baseline and end of study data) in one database. This would have given us the possibility to correct for these differences. Yet, not all researchers responded to our request for additional information.

To investigate changes between participants and changes over time the paired sample T-tests,[65,181,186] Friedman two-way test,[64] the mixed-effects model,[76] and repeated measure analyses[63,65] were used in the included studies. The mixed-effect model and repeated measure analyses are more specific than comparing group mean changes. Therefore, the results of the studies using the better statistical methods are possibly better than the ones using simple statistical techniques. However, in this review we were not able to use this information in the outcome.

Quality of the evidence

By grading the evidence according to the GRADE criteria[183] the overall quality of the studies varied between low and very low. Due to risk of bias, inconsistency, indirectness, imprecision, possible publication bias, or a combination of these, the qualities of the studies were downgraded. None of the articles was eligible for upgrading. The quality of the evidence is summarized in the 'Summary of findings' table (Table 3.1). The small number of participants in the trials was the main reason for the low-quality scores. This is often the case in studies in a pediatric population, and in case of newly introduced interventions. More and larger well-controlled studies are needed to improve the quality and the quantity of evidence. This also emphasizes the need for a core-set of outcome measures in exercise-related research in childhood chronic conditions.[209]

Between the six studies there is a considerable degree of heterogeneity on mode and intensity of the exercise interventions. When assessing heterogeneity the 9-minute-walk-run test, and BMI showed no substantial heterogeneity (respectively $I^2 = 44\%$ and 48%) between the 2 trials. However, the timed up-and-down stairs test, did show substantial heterogeneity ($I^2 = 84\%$).

Potential biases in the review process

The search strategies for MEDLINE/PubMed, EMBASE/OVID, CENTRAL were formulated by the Cochrane Childhood Cancer Group. In addition, two other databases were searched by the use of a search strategy we developed ourselves: CINAHL and PEDro. The PEDro database was difficult to search. Although it is possible that we missed 1 or 2 studies from this database. However, due to the great overlap between results of the different databases it is very unlikely that studies were missed.

Agreements and disagreements with other studies or reviews

In 2010, a review on childhood cancer and physical activity was published by Winter et al. (2010).[210] This review included 28 studies, and almost half had an uncontrolled study design. In eight studies healthy controls were used. Of the four RCTs included in that review, one study included long-term childhood cancer survivors (mean 12 years from diagnosis). Another RCT offered a two to four day intervention, which therefore did not match with the inclusion criteria of this Cochrane review.[86] The two remaining RCTs of the review by Winter et al. (2010) are also included in this Cochrane review.[63,65] A second review on exercise interventions for childhood cancer patients was performed by Huang et al. (2011).[211] They included many of the same studies, but also the study of Chamorro-Vina et al. (2010), which again introduced an intervention of less than four weeks.[80] Both reviews concluded that results are promising, but that there is a need for more and larger RCTs. Both reviews stated that only a subgroup of the childhood cancer population was tested, since almost all studies concerned children with ALL. These findings are consistent with our findings.

The third review on exercise intervention for children with cancer was performed by Wolin et al. (2010).[212] This review studied exercise intervention for adults and/or pediatric cancer patients. They included 12 studies on childhood cancer patients/survivors who did not received a stem cell transplantation and two studies in childhood cancer patients who did receive a hematopoietic stem cell transplantation. Next to the transplanted patients, they included also uncontrolled studies and studies with a short intervention duration.

In the fourth review on this subject was performed by Baumann et al. (2013).[213] This review included 17 studies, of which seven were uncontrolled. They included three RCTs also in our list, but also two others with a short intervention period. Both Wolin et al. (2010) and Baumann et al. (2013) included the study of Shore and shepard (1999) which we excluded for this review because it used healthy controls.[98]

Despite the different studies included in the reviews all the conclusions are comparable with ours: although studies have limitations in their methodology the results are promising. But more research is needed to increase the level of evidence.

AUTHORS' CONCLUSIONS

Implications for practice

Based on the currently available evidence from the included RCTs and CCTs we are not able to draw conclusions regarding the best physical exercise training intervention, neither can we provide information on the best timing of the intervention during or after cancer treatment. However, the six included studies did show that exercise training is feasible in children with ALL.
Effects of the intervention are not yet convincing due to small numbers of participants and insufficient study methodology. Despite that, first results showed more improvements on the outcomes in the intervention group than in the control group. Especially when assessing outcomes such as cardiorespiratory fitness, body composition, flexibility, muscle strength and HrQoL. However, no significant differences were identified for the level of daily activity, fatigue, and adverse events. Moreover, the included studies did not assess activity energy expenditure, time spent exercising, anxiety, depression, or self-efficacy as a study outcome.

Implications for research

The observed heterogeneity in study findings can be due to differences in the physical exercise training intervention (mode, intensity, frequency, duration, as well as location), different outcome measures (quantitative, qualitative, physical, or psychosocial), and methods to assess the effects of an intervention. Consensus on these items is needed in order to facilitate comparison of results across different studies.
More and high-quality evidence is needed in order to be able to draft exercise and physical activity guidelines for this population. We urge the pediatric oncology community to design national or international multi-center studies, while local and small-scale studies must be discouraged.
In addition, since – even in this update – we could only include six RCTs or CCTs with a total of 172 children, there is a need for additional well-designed studies with large sample sizes. Results of ongoing trials have to be awaited, and further trials with adequate power are needed.

Differences between protocol and review

The review is an updated version of the review published in 2013. However, the differences between the review and the protocol remain on a number of aspects.

Instead of using the Cochrane Childhood Cancer Group module for the risk of bias, we used the latest update, which was described in the *Cochrane Handbook for Systematic Reviews of Interventions* of March 2011 to assess the risk of bias of the included studies.[182]

The study of Hartman et al. (2009) included children at diagnosis who were aged one to 18 years. In the protocol we reported our intention to include studies with participants older than three years of age. We opted to change this because some of the studies introduced a tailored exercise program that could be adjusted for the child's age. To see changes in outcomes a child needs to be trainable, co-operative, and testable. For intensive training, which we had in mind when writing the protocol, children aged less than three years will not be able to complete the exercises. However, the study of Hartman et al. (2009) did not assess the effect of a structured intensive training program, but included physiotherapy sessions with exercises that were appropriate for all ages.

We added possible tests that could have been used to assess the primary outcome.

Finally, we added the clinical trial database as resource for the search of ongoing trials (www.clinicaltrials.gov). We also searched the clinical trial database for missed studies.

Table 3.2 Characteristics of studies awaiting classification

Braam 2014	
Methods	Type of study: multi-centre RCT Setting: the Netherlands Department: pediatric oncology Randomization: block-randomization, performed by a blinded independent worker of the pediatric oncology department Stratification: the participants were stratified by (i) cancer (hematological versus solid cancer), (ii) gender and age (boys under 12 vs. ≥ 12 years and girls under 11 years vs. ≥ 11 years) and (iii) during or after cancer treatment Timing: children were during or within the first year following childhood cancer therapy. Children who were during treatment were treated and followed-up in the outpatient clinic, without overnight hospital staying. Study duration: short-term effects are assessed 4 months from baseline. End point measurements: At baseline and after 4 months.
Participants	N = 68 Diagnosis: childhood cancer (treated with chemotherapy, radiotherapy, or both) Age at start study: 8-18 years Sex: 37 boys and 31 girls Exclusion criteria: receiving a bone marrow transplant as a part of the childhood cancer treatment, receiving growth hormones as a part of the childhood cancer treatment, permanent wheelchair use/inability to ride a bike, retardation/inability to make a self-reflexion and follow sports instructions.
Interventions	The 12-week intervention consisted of a combined physical exercise (twice per week) and psychosocial support program (once every 2 weeks). The physical exercise program includes a protocol with both cardiorespiratory and muscle strength training. The sessions are guided by a pediatric physiotherapist and performed at a local (pediatric) physiotherapist institute. The psychosocial support program (6 child and 2 parent sessions) contained psycho-education and cognitive-behavioral therapy (guided by a pediatric psychologist and performed at the academic treatment hospital) The control group received usual-care.
Outcomes	**Primary outcome:** *Cardiorespiratory fitness:* Peak oxygen uptake (ml/kg/min) **Secondary outcome:** *Health-related quality of life:* generic core score, cancer module and multidimensional fatigue module
Notes	This study was published as a conference paper. The total study outcomes were not yet published as a full text article. Not all provided information was presented in the conference paper. We were able to complete the provided information by own source.
Elkateb 2007	
Methods	Type of study: single center CCT Setting: Egypt Department: pediatric oncology Randomization: not performed Stratification: not included Timing: children were during chemotherapy treatment for cancer Study duration: not mentioned End point measurements: at baseline, daily in the first week, after the first week, in the third week and in the sixth week
Participants	n = 50 Diagnosis: childhood cancer Age at start study: preschool- and school-aged children Sex: not mentioned Exclusion criteria: not mentioned

Interventions	Undefined exercise program for the intervention group Undefined program for the control group
Outcomes	**Physical fitness:** *Level of daily activity:* observational checklist for recording activities **Secondary outcomes:** *Fatigue:* observational checklist for sleeping conditions
Notes	This study was published as a conference paper. Based on the currently available information it was not possible to decide if this study was eligible for inclusion in this review

Sabel 2013	
Methods	Type of study: RCT including a cross-over procedure after 10 weeks Setting: Sweden Department: pediatric oncology Randomization: no information in the abstract provided Stratification: no information in the abstract provided Timing: children were 1 to 5 years post diagnosis Study duration: no information provided End point measurements: baseline and after 10 weeks
Participants	n = 13 Diagnosis: Childhood brain tumor survivors (treated with at least radiotherapy); 1 to 5 years post diagnosis Age at start study: 7-17 years Exclusion criteria: not mentioned
Interventions	Intervention: An exercise gaming intervention. The 10-week home-based motion-controlled video console (Nintendo Wii) exercise intervention was performed 30 minutes per day, at least 5 days a week. Children had weekly contact with a coach by videoconferencing. Control intervention: waiting list group
Outcomes	**Primary outcomes:** *Activity energy expenditure:* Energy expenditure and Metabolic Equivalent of Task (METs) assessed by a multisensory activity monitor (SenseWearPro2 Armband)
Notes	This study was published as a conference paper. The total study outcomes were not yet published as a full text article.

Senn-Malashonak 2014	
Methods	Type of study: RCT Setting: Germany Department for Stem Cell Transplantation and Immunology Randomization: no information provided in the abstract Stratification: no information provided in the abstract Timing: children were during hematopoietic stem cell transplantation Study duration: 200 days End point measurements: at hospital admission, at discharge and at 200 days post transplantation
Participants	N = 50 Diagnosis: Childhood cancer patients who received a hematopoietic stem cell transplantation (SCT) Age at start study: not mentioned Sex: not mentioned Exclusion criteria: not mentioned
Interventions	Intervention group: exercise intervention including a standardized resistance, endurance and flexibility training Control group: performed a mental training and relaxation exercises During inpatient treatment the daily sessions last about 40-60 min for each group.

Outcomes	**Primary outcome:**
	Cardiorespiratory fitness: Peak oxygen uptake (ml/kg/min) and 6-Minute-Walking-Test)
	Muscle strength: isometric muscle strength.
	Body composition
	Secondary outcomes:
	Health-related quality of life: questionnaire
Notes	This study was published as a conference paper. The recruitment continued until December 2014

Footnotes: RCT: randomized controlled trial; CCT: controlled clinical trial; ml.kg.min: milliliters per kilogram per minute; VO2max: maximal oxygen uptake.

Table 3.3 Characteristics of ongoing studies

Chamorro-Vina 2012	Exercise in paediatric autologous stem cell transplant patients: a randomized controlled trial
Methods	Type of study: multi-centre RCT
	Setting: Alberta Children's Hospital, Canada
	Department: paediatric oncology and blood and marrow transplantation
	Randomisation: Participant allocation is generated by the software Research Randomizer, a research tool provided by the Social Psychology Network (www.socialpsychology.org). Sealed, non-transparent envelop are used for the actual randomization, which contains a number that allocate the subject to the control or intervention group.The research coordinator is responsible for the random allocation
	Stratification: no information provided
	Timing: children at least four weeks before the hospitalization to start autologous stem cell transplantation when the study information is provided. Study duration: 180 days
	End point measurements: At baseline, 30 days, 85 days and 180 days for general study outcomes. Blood levels at +7, +15 and +56 days post reinfusion.
	Trial register: Gov: NCT01666015
Participants	N = 24
	Diagnosis: all cancer types (common types will be: sarcoma, lymphoma, neuroblastoma, germ cell tumour)
	Inclusion criteria: All childhood cancer patients who undergo autologous stem cell transplantation in Alberta Children's Hospital; who receive myeloablative conditioning; participation approval by treating oncologist; children need to verbally express assent to participate
	Age at start study: 5-18 years
	Exclusion criteria: evidence for cardiac or pulmonary failure associated with treatment (shortening fraction >27%; ejection fraction > 49%); functional or cognitive limitations that would prohibit performance of the home-based training
Interventions	*Intervention group:* An inpatient and outpatient mixed exercise program including both resistance and aerobic training.
	Phase 1. Inpatient phase (in Alberta Children's Hospital during conditioning and the isolated phase of the transplantation). Supervised aerobic training (20-30 minutes) and resistance training of 12-15 repetitions per exercise training, performed 5 times a week.
	Phase 2. the outpatient phase will take place after discharge. The participant will participate in a 10 week mixed supervised and home-based exercise program utilizing the Nintendo® Wii device. Phase 2 includes both supervised exercise training sessions at the University of Calgary (1/week) and home-based training (2/week) including 20-30 min aerobic exercises and 30 minutes of strength and stretching exercises, using the Wii (fit, sports and dance games).
	Children will train with an aerobic exercise intensity between 50-70% of the heart-rate reserve.

	No intervention will take place when: platelet counts <10.000/μl, haemoglobin levels < 8 g/dl, fever >38°C, pain, diarrhoea, haemorrhage or other complications. *Control group:* usual care, waiting list group. During the course of the study these children will not participate in any scheduled exercise program and will perform the same battery of tests as the intervention group. After the end of study (phase 2) they will be offered an exercise program to be held in a separate the University of Calgary.
Outcomes	**Primary outcomes:** *Cardiorespiratory fitness:* submaximal aerobic test on a treadmill. *Muscle endurance:* Partial curl-up, modified push-up test and the sit-to-stand test *Muscle strength:* hand-held dynamometer assessing knee extensions and grip strength *Body composition:* BMI, fat mass estimation by an equation and skin fold measurement *Flexibility:* sit-and-reach test *Activity energy expenditure and level of daily activity:* accelerometer and exercise diary **Secondary outcomes:** *HrQoL:* PedsQL Generic *Fatigue:* PedsQL Multidimensional Fatigue scale
Starting date	June 2012
Contact information	Carolina Chamorro-Vina (phone: 403-210-8482; Email: cchamorro@kin.ucalgary.ca)
Notes	Estimates study completion July 2015
Cox 2011	Physical Activity to Modify Sequelae and Quality of Life in Childhood Acute Lymphoblastic Leukaemia (PAQOL)
Methods	Type of study: single-centre RCT Setting: USA Department: paediatric oncology Randomisation: not described Stratification: not included Timing: children were in the second to eighth day of the ALL treatment protocol Study duration: 135 weeks End point measurements: at baseline (BMD, HrQoL), after 8 weeks (HrQoL), after 15 weeks (HrQoL), and at completion of therapy (BMD and HrQoL) clinicaltrials.gov/ct2/show/NCT00902213
Participants	N = 208 Diagnosis: newly diagnosed with ALL (immunophenotypic diagnosis of non-B cell ALL) Age at start study: 4-18 years Exclusion criteria: age < 4 years or ≥ 19 years at diagnosis, no parents or legal guardian (≥ 18 years) of the study subject who speaks and understands the English language, a diagnosis of cerebral palsy or Down's syndrome, children with a second malignancy, chromosome breakage syndrome, or severe congenital immunodeficiency, inability to obtain written informed consent from parent/young adult and child assent, or females who are pregnant
Interventions	Tailored parent- and child-focused physical activity program An advanced practice nurse will meet twice weekly with the child and family for the first 4 weeks of the intervention to initiate the motivation-based dialogue and therapeutic interaction; this will be followed by once weekly visits during weeks 5-8 of the intervention; and monthly visits during weeks 9 through to end of therapy The physiotherapist will meet at least once weekly with the child and family during weeks 1-4 to initiate the prescriptive tailored exercise program; subsequent visits to reinforce and modify the program will occur at least once every other week during weeks 5-8, and at least once monthly during weeks 9-135 of the intervention. The physiotherapist will visit at least once weekly during weeks 1-4, at least once every other week during weeks 5-8, and at least once monthly during weeks 9-135. During weeks 9-135 of the intervention, the advanced practice nurse will call between the monthly in-person visits to assure fidelity to the intervention and to provide booster support to the intervention where needed

Outcomes	**Physical fitness:** *Muscle endurance/strength:* muscle strength, range of motion, endurance, gross motor skills, used method is not specified *Body composition:* BMD and bone mineral content *Flexibility:* range of motion **Secondary outcomes:** *Health-related quality of life:* method used not mentioned in the protocol *Adverse events*
Starting date	November 2009
Contact information	Cheyl Cox, info@stjude.org
Cox 2011	Physical Activity to Modify Sequelae and Quality of Life in Childhood Acute Lymphoblastic Leukaemia (PAQOL)
Dubnoz-Raz 2010	The Effect of Physical Activity on the Mental and Physical Health of Children With Cancer in Remission
Methods	Type of study: CCT with parallel assignment Setting: Israel; Sheba medical Center Department: no information provided Randomisation: not performed Stratification: not performed Timing: less than 3 months form hospitalization End point measurements: Trial register: Gov: NCT01645436
Participants	N = 22 Diagnosis: childhood cancer Age at start study: 6-16 years *Inclusion criteria:* in remission from cancer; more than 6 months after completion of all therapy *Exclusion criteria:* no written informed consent;refusal of tests: blood, fitness, questionnaires or Dual Energy X-ray Absorptiometry;locomotive handicaps; extreme fatigue, nausea, dyspnoea; concurrent acute illness; recent (less than 3 months) hospitalization; documented (echocardiographic or nuclear medicine) decrease in cardiac function; abnormal blood tests: Hb < 10 gr/dl, neutropenia < 500/mm³, thrombocytopenia < 50,000/ mm³; additional chronic health conditions unrelated to cancer (e.g. celiac disease, cerebral palsy, Down's syndrome)
Interventions	Intervention group: two supervised, 60-minute weekly exercise sessions and instructions to perform additional physical activities throughout the day Control group: usual care; no intervention
Outcomes	**Primary outcomes:** *Cardiorespiratory fitness* *Body composition* and bone mineral density **Secondary outcomes:** *HrQoL* *Anxiety and depression:* Mood
Starting date	February 2010
Contact information	no contacts or location information provided
Notes	Data collection finished in August 2012

Kauhanen 2014	Active video games to promote physical activity, motor performance and quality of life in children with cancer: an intervention study with 2-year follow-up
Methods	Type of study: RCT Setting: Turku University Hospital, Finland Department: paediatric oncology Randomisation: is performed by a computer-generated list - including block randomization with randomly selected block sizes of 2 to 4. Stratification: no information provided Timing: Children are asked to participate within a week after their cancer diagnosis, or as soon as possible after that. End point measurements: baseline; the first week of the intervention, 2 months, 6 months, 1 year and 2.5 years from baseline Trial register: Gov NCT01645436
Participants	N = 40 Diagnosis: Acute Lymphoblastic Leukemia, or other diagnoses outside the central nervous system (e.g. M Hodgkin, non-Hodgkin lymphomas, neuroblastoma, Wilms' tumour, rhabdomyosarcoma, retinoblastoma and Ewing sarcoma) Age at start study: 3-16 years Inclusion criteria: chemotherapy treatment with Vincristine given for childhood cancer in Turku University Hospital or Tampere University Hospital Exclusion criteria: other diseases limiting their in physical and cognitive function; epilepsy, or not able to communicate (in Finnish, Swedish or English)
Interventions	*Intervention group:* The exercise intervention is based on active video gaming on the Nintendo Wii on a light-to-moderate activity level. The intervention includes information and recommendation for physical activity games suitable for performing during the intervention period. Exercises are daily for at least 30 min per day. The 8-week intervention is provided during hospitalization and at home, with considerations of the participants individual conditions. The physiotherapist contacts the participants in the intervention group via telephone for consultation during the intervention aiming to increase participation. Physical activity is not allowed during fever, vomiting or nausea episodes, or if the medical conditions changes. A cardiologist performs regular echocardiograms after anthracycline therapy. *Controlgroup:* Receives general advice for physical activity for 30 min per day and no guidance on playing active video games. In relation to usual care, when needed a physiotherapist is consulted.
Outcomes	**Primary outcomes:** *Body composition:* body mass index, waist circumference *Activity energy expenditure:* a three-dimensional accelerometer; *Level of daily activity:* questionnaire to assess leisure time physical activity in Metabolic equivalents (METs); activity diary; an open question interview about physical activity **Secondary outcomes:** *Fatigue:* PedsQL Multidimensional Fatigue Scale *Adverse events*
Starting date	January 2013
Contact information	MIkko Alola (Tel: not provided, Email: anloka@utu.fi)
Notes	Estimated completion of the primary outcome data December 2016

Mabbott 2013	The neuro-protective effects of exercise in children with brain tumors
Methods	Type of study: RCT Setting: the Hospital for Sick Children, Toronto Canada Department: pediatrics Randomization: no specific information provided Stratification: no information provided Timing: Children finished the cancer treatment End point measurements: Baseline, week 26-29 and week 42-45 Trial register: Gov: NCT01645436
Participants	N = 30 Diagnosis: childhood brain tumor survivors Age at start study: 9-14 years Inclusion criteria: Native English speaker or at least 2 years of schooling in English at time of the inclusion; diagnosed with a hemispheric or posterior fossa tumor and treated with cranial spinal or focal radiation Exclusion criteria: more than 7 years post diagnosis; have a prior history of traumatic brain injury; neurologic disorder, visual or sensory impairment, cerebral palsy, developmental delay or learning disability; requiring sedation for MRI imaging; severe neurological/ motor dysfunction that would preclude safe participation in an exercise program
Interventions	*Intervention group:* receive a 12-week exercise intervention, with 90 min sessions provided 3 times a week. Each session includes 30 min of circuit training, 30 min of organized sports (soccer, floor hockey, basketball, etc.) and 30 min of socializing including a healthy snack *Control group:* receives a delayed intervention. They receive the same 12-week intervention after the intervention group completed the intervention period.
Outcomes	**Primary outcomes:** *Cardiorespiratory fitness:* Physical fitness: VO_{2max}
Starting date	July 2013
Contact information	Donald Mabbott (Tel: 416-813-8875; email: donald.mabbott@sickkids.ca)
Soares-Miranda 2013	Physical activity in pediatric cancer patients with solid tumors (PAPEC): trial rationale and design
Methods	Type of study: RCT Setting: Childen's Hospital of Madrid, Hospital Infantil Universitario Nino Jesus, (Spain) Department: no information provided Randomization: no specific information provided Stratification: on sex Timing: during neo-adjuvant chemotherapy End point measurements: baseline, middle measurement during the intervention period, end of the intervention, 2 months after the intervention Trial register: Gov: NCT01645436
Participants	N = 60 Diagnosis: pediatric patients with extra-cranial primary solid tumors during a course of neo-adjuvant chemotherapy Age at start study: 4-18 years Inclusion criteria: new diagnosis of an extra-cranial solid tumor; having a good performance status; receiving treatment in Hospital Infantil Universitario Nino Jesus of Madrid, living in the Madrid province Exclusion criteria: previously receiving cancer therapy
Interventions	Intervention group: receives a 60-70 min supervised in-hospital combined exercise-training program (aerobic and strength) 3/week. The intervention includes both 30 min aerobic and 30 min strength training. Muscle strength training is performed with dumbbells and pediatric weight machines; performing 2-3 sets of 8-15 repetitions.

	The intervention is offered between the whole period of neoadjuvant chemotherapy treatment. Related to the tumor type, the period can range between 4-24 weeks. The intervention is tailored to the patient according to training guidelines and/or performed in the hospital room when needed. Control group: receives usual hospital care; no scheduled training but physical therapy when needed and recommendations for a healthy lifestyle.
Outcomes	**Primary outcomes:** *Cardiorespiratory fitness:* Cardio-respiratory capacity: VO_{2peak} *Muscle endurance/ strength:* 6 repetition maximum of leg and chest (bench) presses and lateral rowing, timed up and down stairs test, 3m and 10m timed up and go test *Activity energy expenditure:* uni-axial accelerometer **Secondary outcomes:** *HrQoL:* the Child report form of the Child's Health and illness Profile-Child Edition (CHIP-CE/CRF), adolescent edition (CHP-PE/AE) and parents edition (CHP-CE/PRF)
Starting date	September 2012
Contact information	L. Soares-Miranda (Tel: +351962591421; Email: soaresmiranda@fade.up.pt;
Notes	Estimated end date September 2015
Wiskemann 2012	Physical activity intervention program for childhood cancer patients under chemo- and /or radiation therapy
Methods	Type of study: CCT Setting: Hospital of the University Hospital of Heidelberg, Germany Department: no information provided Randomization: no randomization Stratification: no stratification Timing: less than 8 weeks after childhood cancer diagnosis End point measurements: at baseline, 3, 6, 9 and 12 months, questionnaires at 6 and 12 months Trial register: Gov: NCT02216604
Participants	N = 60 Diagnosis: primary pediatric cancer diagnosis (leukemia, brain and bone tumors) Age at start study: 5-21 years Inclusion criteria: date of diagnosis no longer than 8 weeks before start of the study Exclusion criteria: sever cardiac impairment; bone metastasis inducing skeletal fragility; other orthopedic diseases or any other circumstances that would impede ability to give informed consent or adherence to study requirements
Interventions	Intervention group: Multimodal exercise intervention. The intervention includes a 3-5 weekly guided training program of 15-30 min sessions based on endurance, strength and balance game training using the Wii (Nintendo) as well as on age-specific resistance training and sessions of body awareness. During the outpatient phase of the treatment there is a home-based exercise training (3-5 weekly) using a manual. The patients also obtain a movement diary and pedometer. Control group: age, disease and gender matched group receiving no intervention
Outcomes	**Primary outcome:** *Muscle endurance/strength:* hand-held dynamometer, Timed Up-and-Down Stairs test, one-leg stand, posturomed and force plate *Body composition* *Flexibility:* goniometer **Secondary outcome:** *HrQoL*
Starting date	December 2012
Contact information	Joachim Wiskemann (phone: +49-6221-565904; email: joachim@wiskemann-online. de or andrea.kulozik@med.uni.heidelberg.de)

Footnotes: ALL: acute lymphoblastic leukemia; BMD: bone mineral density; BMI: body mass index; CPET: cardiopulmonary exercise test; DXA: dual-energy x-ray absorptiometry; HrQoL: health-related quality of life; RCT: randomized controlled trial; SRT: steep ramp test; $VO_{2\,peak}$: maximal oxygen consumption.

Table 3.4 Characteristics of included studies

Hartman 2009	
Methods	Design: single center RCT Setting: the Netherlands Department: pediatric oncology/hematology, pediatric physiotherapy, pediatric endocrinology Randomization: blinded for investigators and treating physicians Stratification: not mentioned Study duration: 3 years. Duration of the intervention: 24 months. Follow-up duration: 12 months Timing: inclusion started directly after diagnosis, at the beginning of their chemotherapy treatment End point measurements: at diagnosis, 32 weeks after diagnosis, 1 year after diagnosis, at the end of treatment (and 2 years after diagnosis), 1 year after the end of treatment. Six weeks following diagnosis there was 1 additional measurement
Participants	n = 51 Diagnosis: ALL (ALL non-high risk n = 34, ALL high risk n = 17) Age at start study: median age: 5.4 years (range 1.3 to 17.1 years) Sex: 30 boys, 21 girls Exclusion criteria: children with low cognitive impairment and those which could not understand the Dutch language
Interventions	The intervention consisted of an exercise program of 2 years. The program consisted of a hospital-based program performed by pediatric physiotherapists. During these sessions, the physiotherapist measured the motor function to ensure an optimal level of motor functioning. In addition, there was a home-based exercise program. Parents were supplied with an exercise list, enabling them to select exercises most appropriate for their child's age and also to vary exercises. The exercise program included exercises to maintain ankle dorsiflexion mobility and short-burst high-intensity exercises, to prevent reduction of BMD. In addition, there were exercises to maintain hand and leg function. The hand and leg function exercises were performed once a day; stretching and jumping exercises twice daily. The duration of an exercise session was not mentioned When necessary the exercise program was adjusted during these sessions The control group received care as usual
Outcomes	**Physical fitness:** *Body composition:* BMI, lean body mass, and percentage body fat. The lean body mass and body fat were measured by DXA (lumbar spine and total body) *Flexibility:* passive ankle dorsiflexion; the range of motion past the neutral position received a positive notation and less than neutral a negative notation. Motor performance of children less than 3.5 years of age was assessed by the use of the Dutch BSID-II; ≥ 4 years old by the use of the Dutch version of the Movement-ABC **Secondary outcomes:** None of the secondary outcomes were assessed.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "At diagnosis randomization into the intervention or the control group was carried out in randomly permuted blocks of randomly chosen size, using sealed envelopes prepared by the statistician"
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement of 'low risk' or 'high risk'. The use of assignment envelopes are described, but it remains unclear whether envelopes were sequentially numbered, or opaque

Blinding of participants and personnel (performance bias)	High risk	Participants and parents were not blinded for randomization; this was unclear for physiotherapists. The investigators and treating physicians were blinded for the study randomization
Blinding of outcome assessment (detection bias)	Low risk	Outcome assessors who performed the study outcome tests were blinded for study randomization
Incomplete outcome data (attrition bias)	High risk	The study authors used a simple imputation method: for children who did not complete the study, data prior to elimination were included. No further information was provided on the imputation of some value for missing data
Selective reporting (reporting bias)	Low risk	All primary and secondary outcome measures were listed in the methods section and reported in the results section
Other bias	Low risk	There was no baseline imbalance found, the baseline differences between both groups were not significant. In addition, the number of measurements did not differ for the intervention group or control group

Macedo 2010

Methods	Design: single center RCT Setting: Brazil Department: pediatric oncology/hematology Randomization: random assignment but no further specifications available Stratification: not mentioned Study duration: 10 weeks Timing: inclusion of the study started during maintenance therapy of the childhood ALL treatment End point measurements: in the intervention group at baseline plus an evaluation every alternate week. In the control group at baseline, and 10 weeks thereafter
Participants	n = 14 Diagnosis: ALL Age at start study: mean age of the whole group was 8.3 ± 2.6 years (range 5 to 14 years). The mean age of the intervention group was 7.0 years and that of the control group 9.0 years Sex: 5 boys and 9 girls Exclusion criteria: children with a chronic lung disease, neuromuscular disease, or those receiving or having received radiotherapy treatment
Interventions	This study investigated an inspiratory muscle training program. They studied the effects of a domiciliary inspiratory muscle training with a duration of 15 minutes, performed twice a day, for 10 weeks. The training was performed with a threshold device using a load of 30% of the maximal inspiratory pressure. The control group received care as usual
Outcomes	**Physical fitness:** *Muscle endurance/strength:* respiratory muscle strength (maximal inspiratory pressure and maximal expiratory pressure) assessed with a digital manometer **Secondary outcomes:** None of the secondary outcomes were assessed
	Article was written in Portuguese

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Children were randomly selected and randomly assigned to 2 groups, but the exact randomization methods were not reported

Allocation conceal-ment (selection bias)	Unclear risk	The exact randomization methods were not reported. It was not clear whether the researchers used sealed envelopes, central allocation, or another method
Blinding of partici-pants and person-nel (performance bias)	High risk	The study did not address the blinding of participants and personnel. However, due to the nature of the interventions blinding was virtually impossible
Blinding of out-come assessment (detection bias)	Unclear risk	The study did not address blinding of outcome assessment
Incomplete out-come data (attri-tion bias)	Unclear risk	Insufficient reporting: the authors stated that sample losses occurred; however, they did not report the reasons for these sample losses, neither did they provide information on the used imputation methods
Selective reporting (reporting bias)	Low risk	Respiratory muscle strength was the primary outcome. By assessing and reporting on (changes over time of) both the maximal inspiratory pressure and maximal expiratory pres-sure there was no selective reporting of the study data
Other bias	High risk	Differential diagnostic activity: the intervention group and the control group received an unequal number of measure-ments. However, this study was free of baseline imbalance; the base-line differences between the control group and intervention group on outcome related items were not significant

Marchese 2004

Methods	Type of study: single center RCT Setting: USA Department: pediatric rehabilitation, pediatric oncology, pediatric physiotherapy Randomization: primary investigator offered the children an envelope to select assignment into the intervention or control group Stratification: children were stratified according to their childhood cancer risk group and first versus second part of the maintenance therapy Study duration: 4 months Timing: inclusion of the study started during maintenance therapy End point measurements: at baseline and 4 months later
Participants	n = 28 Diagnosis: ALL Age at start study: median age of the whole group was 7.7 years (range 4.3-15.8 years). The median age of the intervention group was 7.6 years (range 4.3-10.6 years) and of the control group 8.6 years (range 5.1-15.8 years) Sex: 20 boys and 8 girls Exclusion criteria: a history of antecedent neurological, developmental, or genetic disorders and those receiving a physiotherapy intervention at the start of the study
Interventions	The intervention program included 5 hospital-based physiotherapy sessions (week 0, 2, 4, 8, and 12) of 20-60 minutes. The first session was performed immediately after the baseline testing Next to the hospital-based program, the program also included an individualized home exercise program. This program consisted of ankle dorsiflexion stretching exercises (30 seconds, 5 days a week), bilateral lower extremity strengthening exer-cises (3 sets of 10 repetitions, 3 days a week), and aerobic exercise (daily). The aerobic exercise could be walking, cycling, or swimming; chosen by the participant The control group received care as usual

Outcomes	**Physical fitness:** *Cardiorespiratory fitness or peak work rate:* 9-minute run-walk test and the timed up-and-down stairs test *Muscle endurance/strength:* knee extension strength and ankle dorsiflexion strength both tested with a hand-held dynamometer. This study also used the time up-and-down stairs test and the 9-minute run-walk test *Flexibility:* ankle dorsiflexion range of motion **Secondary outcomes:** *Health-related quality of life:* PedsQL version 3.0 *Adverse events:* any negative effect from the exercises or experienced complications attributed to the physical program

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The children were stratified by risk group and by whether they were in the first or second half of the maintenance therapy. After that the primary investigator offered the children an envelope to select assignment into the intervention or control group
Allocation concealment (selection bias)	Unclear risk	Insufficient information to permit judgement of 'low risk' or 'high risk'. The use of assignment envelopes is described, but it remains unclear whether envelopes were sealed, sequentially numbered, or maybe opaque
Blinding of participants and personnel (performance bias)	High risk	Participants and parents were not blinded for randomization; for personnel this was unclear
Blinding of outcome assessment (detection bias)	Low risk	The outcome assessors for hand-held dynamometry , the timed up-and-down stairs test and the 9 minute run-walk test were blinded for study randomization. Therefore these items had a low risk for detection bias The PedsQL (quality of life) questionnaires were filled in by both parents and children. Parents and children were not blinded for the study randomization and therefore the quality of life assessment was found to be of high risk for detection bias We judged the overall risk of detection bias for this item to be low because the researchers blinded outcome assessors as much as possible
Incomplete outcome data (attrition bias)	Unclear risk	The authors reported missing data for daily logs of activity and heart monitor. But no information was reported on methods used for data imputation in case of missing data
Selective reporting (reporting bias)	Low risk	All the pre-specified primary and secondary outcomes of the study were listed in the methods section and reported in the results section
Other bias	Unclear risk	The non-significant baseline differences were reported for patient characteristics, however, not for study outcome measures. It remains unclear whether the mean differences between the control group and the intervention group at baseline were significant or not Furthermore we checked for differential diagnostic activity. During the study all children were pre tested and post-tested. The number of measurements did not differ for the intervention group or control group

Moyer-Mileur 2009	
Methods	Type of study: single center RCT Setting: USA Department: pediatric oncology Randomization: not mentioned Stratification: not mentioned Study duration: 12 months Timing: the inclusion of the study started during the ALL maintenance chemotherapy End point measurements: measures of physical size were obtained at baseline and every 3 months, physical activity was measured at baseline and at 6 and 12 months
Participants	n = 14 Diagnosis: standard-risk ALL Age at start study: mean age (± SD) of the intervention group was 7.2 ± 0.7 years and the mean age of the control group was 5.9 ± 0.7 years Sex: 7 boys and 6 girls; 1 unknown (drop-out) Exclusion criteria: not mentioned
Interventions	The intervention included a 12-month home-based exercise and nutrition program Children were prescribed to perform a minimum of 3 'fifteen to twenty-minute' sessions of moderate-to-vigorous activity per week. Activity examples were provided on the pyramid for youth and parents were asked to record the type and amount of physical activity, immediately after the activity was performed Children received nutrition education materials on the basis of the United States Department of Agriculture Food Guide Pyramid and nutrition-related activities monthly The control group received care as usual
Outcomes	**Physical fitness:** *Cardiorespiratory fitness or peak work rate:* progressive aerobic cardiovascular endurance run *Muscle endurance/strength:* push-ups, the sit-and-reach test *Body composition:* BMI, muscle mass (measured by the analysis of the tibia using peripheral quantitative computed tomography *Flexibility:* sit-and-reach distance test *Level of daily activity:* pedometer combined with an activity diary (monthly, 2 weekdays and 1 weekend day) and the ACTIVITY GRAM questionnaire **Secondary outcomes:** None of the secondary outcomes were assessed

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The method of randomization was not provided in the article
Allocation concealment (selection bias)	Unclear risk	The method of randomization was not provided in the article
Blinding of participants and personnel (performance bias)	High risk	The study did not address this item. However, due to the nature of the interventions blinding was virtually impossible
Blinding of outcome assessment (detection bias)	Unclear risk	The study did not address this item

Incomplete out-come data (attri-tion bias)	Unclear risk	Although authors reported that 1 child withdraw after 3 months (caused by lack of interest and data of this child were not taken into analysis), the information provided was insuf-ficient to decide whether there this withdrawal could have had influence on the study outcomes
Selective reporting (reporting bias)	Low risk	The article presented both the mean (plus confidence interval or SD) of all outcome variables and figures including the individual changes of the participants
Other bias	Low risk	There was no baseline imbalance found, the baseline differ-ences between both groups were not significant Furthermore we checked for differential diagnostic activity. During the study all children were pre tested and post-tested. The number of measurements did not differ for the interven-tion group or control group

Tanir 2013		
Methods	Type of study: RCT in two university hospitals Setting: Turkey Department: pediatric oncology Randomization: not mentioned Stratification: not mentioned Study duration: 3 months Timing: Being in remission (having received a diagnosis of ALL at least 1 year before the study End point measurement: at 3 months	
Participants	n = 40 Diagnosis: ALL Age at start study: mean age (± SD) of the intervention group was 10.21 ± 1.51 years and the mean age of the control group was 10.72 ± 1.52 years Sex: 24 boys and 16 girls Inclusion criteria: (a) being in the age group 8–12, (b) being in remission (having received a diagnosis of ALL at least 1 year before the study),(c) being followed up as an outpatient, (d) not having participated in a regular exercise program in the last 6 months, (e) residing in Istanbul, (f) displaying a Grade 3 in motor strength (Hislop & Montgom-ery, 2006), (g) not having a previous history of cancer, (h) having no history of neuro-logical disease or genetic disorder before the diagnosis of ALL, (i) having no cardiac, pulmonary, renal or hepatic dysfunction, and (j) having no problem with sight, hearing or perception.	
Interventions	The children in the trial group were offered their first session of training at a desig-nated room in the hospital. One of each child's parents was admitted into the ses-sion to serve as a supporting and motivating force. In the session, the exercises that the children would be doing in the next 3 months were demonstrated. The workout comprised active range of motion (ROM), leg muscle strengthening and aerobic exercises. 1) Active ROM exercises; 5 days a week, 3 times a day, 20 times each repetition 2) Leg Exercises for strengthening the muscles; 3 days a week, 3 times a day. 3) Aerobic Exercises, three times a week, once a day, for a half-hour. The control froup received care as usual (see Notes for additional information).	
Outcomes	**Physical fitness:** *Cardiorespiratory fitness or peak work rate:* 9-minute run-walk test, Timed up-and-down stairs test, *Muscle endurance/strength:* Leg and back strength tested with a dynamometer. Timed Up and Go Test. This study also used the time up-and-down stairs test and the 9-minute run-walk test *Flexibility:* Range of motion; joint unclear **Secondary outcomes:** *Health-related quality of life:* Peds Ql 3.0 Cancer Module Children's Form	

Notes	In the method section of the article authors reported "The children and parents in the control group were given exercise pamphlets after the monitoring and then pro-vided with 30–60 minute training sessions." Nonetheless, they also reported "No exercise was recommended to the patients in the control group over the course of the study." We repeatedly contacted the authors for additional information, but we unfortunately, did not receive a response to our requests. Without additional information and based on the additional information in the paper, we concluded that the authors made a writing-mistake in the method section of their paper. Therefore we interpreted the results as if the control group received usual care.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "The two groups were formed by randomized selection." The method used during the randomization is unclear. .
Allocation concealment (selection bias)	Unclear risk	Quote: "The two groups were formed by randomized selection." The method used during the randomization is unclear.
Blinding of participants and personnel (performance bias)	High risk	Outcome group: Quote: "Participants were asked to prevent for interaction and not speak about the intervention with children/parents of the control group". However: "it was observed that the children and their families who were subjects of the study did interact, forming loyalties and relationships in a social environment. For this reason, when it is considered that the children and families who received the exercise training had the opportunity to share their knowledge and practice with the control group, we believed that our findings might be the result of this interaction process." Personnel: no information about blinding personnel
Blinding of outcome assessment (detection bias)	Unclear risk	The study did not address this item.
Incomplete outcome data (attrition bias)	Unclear risk	Information on all outcomes was provided, yet the authors provided no information about missing data. The authors reported that one child demised in the intervention group. However, data on that particular child were not provided.
Selective reporting (reporting bias)	High risk	Data of the PedsQl 4.0 is only reported as differences between boys and girls and not as differences between the intervention and control group.
Other bias	High risk	Baseline differences in sex, treatment anxiety, and on the Goniometer. Furthermore we checked for differential diagnostic activity. The number of measurements did not differ for the intervention group or control group indicating that this study is free of differential diagnostic activity.

Yeh 2011	
Methods	Type of study: single center CCT feasibility study (quasi-experimental) Setting: Taiwan Department: pediatric oncology Randomization: not performed Stratification: the intervention group and controls were matched by age and sex Timing: the inclusion of the study started during the ALL maintenance chemotherapy (1 week after completion of the dexamethasone treatment) Study duration: 10 weeks End point measurements: at baseline, once weekly during the 5-week intervention, at the end of the intervention and 1 month after the intervention
Participants	n = 24 Diagnosis: ALL Age at start study: mean age intervention group 11.0 ± 3.56 years, mean age of the control group 12.5 ± 3.86 years Sex: 12 boys and 10 girls; 2 unknown (drop-outs) Exclusion criteria: children who were unwilling to perform an aerobic exercise, or those with physical and developmental impairment
Interventions	The intervention consisted of a home-based aerobic exercise instructed by video. 1 session included a warm-up of 5 minutes, aerobic exercise of 25 minutes and a cooling down period of 5 minutes. The exercises were performed at least 3 times a week, over a total of 6 weeks. In addition, children recorded their physical activity and heart rate data during the exercises in a physical activity log for 3 days with 24 1-hour blocks The aerobic exercise sessions aimed to increase 40-60% of the child's heart rate reserve The control group received care as usual
Outcomes	**Physical fitness:** None of the physical fitness outcomes were assessed **Secondary outcomes:** *Fatigue:* PedsQL multidimensional fatigue scale

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	The researcher-team used a quasi-experimental design that had no random assignment
Allocation concealment (selection bias)	High risk	The researcher-team used a quasi-experimental design that had no random assignment
Blinding of participants and personnel (performance bias)	High risk	The study did not address this item. However, due to the nature of the interventions blinding was virtually impossible
Blinding of outcome assessment (detection bias)	Unclear risk	The study did not address this item
Incomplete outcome data (attrition bias)	Low risk	2 types of analyses were conducted: ITT analysis used the data of all children, and the per-protocol analysis, which included only those children who adhered to the exercise prescription
Selective reporting (reporting bias)	High risk	Not all the pre-specified primary outcomes have been reported. In addition, adherence was mentioned to be extra or a secondary outcome. However, in the results the authors focused in this item

| Other bias | Unclear risk | The non-significant baseline differences were reported for fatigue study outcomes. However, it remains unclear whether the intervention and control group had different baseline scores on the other study outcomes: physical activity log, OMNI walk/run scale, and the stages of change.
Furthermore we checked for differential diagnostic activity. The number of measurements did not differ for the intervention group or control group. Therefore this study was free from differential diagnostic activity |

Footnotes: ALL: acute lymphoblastic leukemia; BMD: bone mineral density; BMI: body mass index; BSID-II: Bayley Scales of Infant development; CCT: controlled clinical trial; DXA: dual energy x-ray absorptiometry; ITT: intention to treat; Movement-ABC: Movement Assessment Battery for Children; OMNI walk/run scale: Omnibus - walk/run scale; PedsQL: Pediatric Quality of Life Inventory; RCT: randomized controlled trial.

Table 3.5 Characteristics of excluded studies

	Reason for exclusion
Chamorro-Vina 2010	The exercise intervention was offered less than 4 weeks
Chung 2014	The exercise intervention was offered less than 4 weeks
Emanuelson 2014	Used a different outcome assessment tool than described in the inclusion criteria
Geyer 2011	No exercise intervention
Gohar 2011	Uncontrolled study
Herbinet 2014	The exercise intervention was offered less than 4 weeks
Hinds 2007	The exercise intervention was offered less than 4 weeks
Huang 2014	No exercise intervention
Jarden 2013	Study assessed in adult cancer patients and the study was an uncontrolled study
Kurt 2011	No exercise intervention
Oldervoll 2011	Study assessed in adult cancer patients
Rief 2011	Study assessed in adult cancer patients
Rosenhagen 2011	case-control study
Ruiz 2010	Uncontrolled study
Shore 1999	Used health normal volunteer children as a control population
Speyer 2010	The exercise intervention was offered less than 4 weeks and cross-over randomized trial without data presentation after the first intervention period (before cross-over)
Speyer 2011	The exercise intervention was offered less than 4 weeks
Steel 2011	Study assessed in adult cancer patients
Te Winkel 2008	This study presents pilot data of a study that was reported by Hartman et al (2009). Hartman et al. was already included in the review (Hartman 2009)
William Li 2013	The exercise intervention was offered less than 4 weeks
Winter 2013	The exercise intervention was offered less than 4 weeks