Exergaming improves balance in children with spastic cerebral palsy
with low balance performance: results from a multicenter controlled trial
Peer-reviewed author version

Exergaming improves balance in children with spastic cerebral palsy with low balance performance; results from a multicenter controlled trial

Pieter Meyns[1*], Ian Blanckaert[2], Chloé Bras[3], Nina Jacobs[1], Jaap Harlaar[3,4], Laura van de Pol[5], Frank Plasschaert[6,7], Hilde Van Waelvelde[2‡], Annemieke I. Buizer[3,8‡]

[1] *Rehabilitation Research (REVAL), Rehabilitation Sciences, Hasselt University, Hasselt, Belgium*

[2] *Pediatric Rehabilitation, Rehabilitation Sciences, Ghent University, Ghent, Belgium*

[3] *Amsterdam Movement Sciences, Rehabilitation Medicine, Amsterdam UMC locatie VUmc, Amsterdam, the Netherlands*

[4] *Biomechatronics & Human-Machine Control, Biomechanical Engineering, Delft University of Technology, Delft, the Netherlands*

[5] *Child Neurology, Amsterdam UMC locatie VUmc, Amsterdam, the Netherlands*

[6] *Cerebral Palsy reference center, University Hospital Ghent, Ghent, Belgium*

[7] *Human Structure and Repair, Medicine and Health Sciences, Ghent University, Ghent, Belgium*

[8] *Emma Children's Hospital, Amsterdam UMC locatie VUmc, Amsterdam, the Netherlands*

[‡] *H. Van Waelvelde and A.I. Buizer are joint senior authors*

*Corresponding author: Pieter Meyns, Rehabilitation Research (REVAL), Rehabilitation Sciences, Hasselt University campus Diepenbeek, Agoralaan gebouw D, B-3590, Hasselt, Belgium. E-mail: Pieter.meyns@gmail.com

- Exergame-training (training using video-games) shows mixed results in children with cerebral palsy (CP).
- Children with spastic CP (GMFCS level I-II) with a high baseline balance-level did not show functional balance improvements after this home-based exergame-training, suggesting that these children should not be enrolled in this type of exergame-training protocol.
- Children with spastic CP (GMFCS level I-II) with a low baseline balance-level showed clinically relevant functional balance improvements after this home-based exergame-training, suggesting that these children can benefit from enrolment in this type of exergame-training protocol to improve their balance.
- The distribution of CP-symptoms did not affect the effectiveness of this balance exergame-training in children with spastic CP with GMFCS-level I and II.

Exergaming improves balance in children with spastic cerebral palsy with low balance performance; results from a multicenter controlled trial

Purpose: Previous studies investigating the effectiveness of exergame balance-training (using video-games) in children with cerebral palsy (CP) yielded inconsistent results that could be related to underpowered studies. Therefore, in this multicenter intervention study, we investigated whether exergaming improves balance clinically in spastic CP.

Materials and methods: In total, 35 children with unilateral or bilateral spastic CP (GMFCS-level I-II) were included (age-range:7-16 years); 16 at VUMC (trial:NTR6034), 19 at UHG (trial:NCT03219112). All participants received care as usual. The intervention group (n=24) additionally performed exergame-training; 6-8 weeks home-based X-box One Kinect training focused on balance. Balance performance was assessed with the pediatric balance scale (PBS) and two subscales of the Bruininks-Oseretsky Test of Motor Proficiency-2nd edition ('balance'[BOTbal] and 'running speed & agility'[BOTrsa]). Mixed model ANOVAs with between and within factors were used to test differences between and within groups.

Results: On group level, no post-intervention differences were found between the intervention and control group (PBS:$p=0.248$,$\eta_p^2=0.040$; BOTbal:$p=0.374$,$\eta_p^2=0.024$; BOTrsa:$p=0.841$,$\eta_p^2=0.001$). Distribution of CP-symptoms (unilateral versus bilateral) did not affect training (PBS:$p=0.373$,$\eta_p^2=0.036$; BOTbal:$p=0.127$,$\eta_p^2=0.103$; BOTrsa:$p=0.474$,$\eta_p^2=0.024$). Children with low baseline balance performance (based on PBS) in the intervention group showed improvements in balance performance after training (PBS:$p=0.003$,$\eta_p^2=0.304$; BOTbal:$p=0.008$,$\eta_p^2=0.258$), whereas children with high baseline balance performance did not.

Conclusions: This exergame-training resulted in balance improvements for the current population of CP that had a low baseline function.

Keywords: cerebral palsy; postural balance; video games; rehabilitation

Introduction

The prevalence of cerebral palsy (CP) is 1.5-3 in 1000 live births in Europe[1], making CP the most common developmental disorder associated with lifelong movement and posture disability[2]. Even though CP is a heterogeneous group with respect to affected brain regions and symptom typology, instability during standing[3] and gait[4] are common major impairments which severely affect the child's independence and risk of falls. As such, balance training is a key aspect in their rehabilitation[5].

Training using commercial virtual reality games (i.e. exergaming) is a relatively new development in rehabilitation, which is increasingly used to train balance in CP[6]. Exergames are deliberately developed to increase therapy compliance or therapy time by focusing on joy and motivational aspects[7]. They are a powerful medium to provide feedback about a person's movements, thus can motivate the person to adapt their movement strategy[8] or to perform more repetitions. As such, the advantages of exergames are based on key concepts of motor learning: motivation, feedback and repetition[9]. Given the potential of exergames for rehabilitation, the effectiveness of several exergames to train balance have been investigated in CP. Many of these studies indicate promising effects[10–14], while others do not[15–17]. The most common limitation in most studies is a low sample size of less than 20 participants in the intervention group[12–17], indicating that the described results may be biased due to the limited and possibly not representative population. Several reviews and meta-analyses underline the promising effects of exergaming to train balance in CP but clearly indicate the need for more studies with higher sample sizes[6,18,19].

Previous studies included different distributions of CP-symptoms to examine the effect of exergaming on balance; some included only children with bilateral CP[10,17]

or unilateral CP[13,16], while some combined the data of the two groups[12,14,15], with varying degree of affected functional mobility. This variability in topographical distribution of CP-symptoms is one factor that may explain the mixed results.

Besides the distribution of symptoms, another factor may affect the results of intervention in CP. It has been suggested that baseline performance can affect the effectiveness of training. Previous research on the effectiveness of sitting balance training using exergames in children with CP who were hospitalized after lower limb surgery has indicated that children with a low performance at baseline can benefit more from training than children that have a high performance at baseline[14]. Previous research on other intervention paradigms (e.g. upper limb function) in children with CP support this finding[20,21].

Therefore, our aim is to investigate whether this exergame balance-training improves balance in children with unilateral and bilateral spastic CP. We hypothesize that children with CP who receive exergame-training will improve balance performance more than those who do not (hypothesis 1), especially children with bilateral CP as both the lower limbs are affected while children with unilateral CP have one side that is less affected which can compensate for the affected side (hypothesis 2). Further, we hypothesize that children with low baseline balance performance will improve more than children with high baseline balance performance (hypothesis 3).

Materials and methods

Participants

In this multicenter cohort intervention study, we combined data from two non-randomized controlled trials (registered trials; NTR6034 & NCT03219112). Parents and children were approached by their physician to ask for participation in this study if they met the following inclusion criteria; (1)7-16 years, (2)spastic CP (unilateral or bilateral), (3)gross motor function classification system (GMFCS) level I–II, (4)no Botulinum-Neurotoxin A injections within 6 months or lower limb surgery within 12 months before baseline tests. Participants were recruited at the Amsterdam University medical center location VUmc (VUMC) and the University Hospital Ghent (UHG). Experiments were approved by the local ethics committee (Medical ethical review commission VUMC - NL57227.029.16 & Commission for medical ethics UHG – B670201525057) and performed with informed written consent from the parents of the participants and children from 12 years and older. Patient recruitment and data collection started in February 2017 (UHG) and January 2017 (VUMC). The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008. Children were divided by the investigators into an intervention and control group according to the availability of the Xbox one consoles and the date of visit of check-up in the hospital. This design was chosen as inclusion was started before the consoles were delivered; i.e. children were only allocated to the control group when the consoles were not yet available, and as soon as consoles were available children were allocated to the intervention group. All children with CP who participated in the study were provided a console after the end of the project. Depending on the site (VUMC or UHG), participants also have performed other assessments (e.g. gait analysis), which

may have influenced the participant's motivation or fatigue. Investigators, participants and parents were not blinded to group allocation.

Intervention

The intervention and control group received care as usual. Additionally, the intervention group performed a home-based balance exergame-intervention using the commercially available X-Box One and Kinect (Microsoft Corporation, Redmond, WA, USA). All children used a TV screen to display the game. They were asked to exercise using specific games 1 session/day, 5 days/week for 30 minutes/session for 6-8 weeks. Children were allowed to choose to play the tennis, soccer and bowling subgames of Kinect Sports Rivals (Rare Ltd., Twycross, UK), as the other subgames required movements that were less directly related to balance (i.e. climbing, shooting and jet ski required only arm movements). Children used their body movements to trigger and play the game. Tennis required arm swing and trunk rotation, soccer demanded standing on one leg, and bowling required swinging of the arm while doing a lunge. Children (and their parents) were instructed to allow for sufficient space so that the participant could move freely in the room. As children with GMFCS level II were included, the use of AFOs was allowed during the exergame sessions for all participants. They were asked to play the games while standing and to have 30 minutes of actual play time per session (excluding time for loading the game). Adequate balance control was required throughout the game plays. Time played on the X-Box on the children's personal account was automatically saved and was provided to the researchers to estimate whether the imposed training duration (minimum 15 hours) was achieved.

Outcome measures

Balance was assessed with several clinical balance scales before and after training; (i) the Pediatric Balance Scale (PBS) is a valid scale to measure functional balance performance in the context of everyday tasks, such as; sit-to-stand, standing balance, reaching forward in children with CP[22]. It has 14 items to score on a 5-point ordinal scale, with a maximum score of 56. (ii) the 'Balance' subtest of the Bruininks-Oseretsky Test of Motor Proficiency-2 (BOTbal) measures control and coordination to maintain balance while standing and walking (e.g. stand with eyes open and closed, stand on one leg, walk on a line)[23]. It is a 9-item scale on a 5-6-point ordinal scale, with a maximum score of 37. (iii) the 'Running Speed and Agility' subtest of the Bruininks-Oseretsky Test of Motor Proficiency-2 (BOTrsa), although originally designed to present fitness function, it also measures gross motor function including balance related stepping (e.g. shuttle run, stepping sideways over a balance beam, one-legged stationary hop)[23]. It is a 5-item scale on a 11-13-point ordinal scale, with a maximum score of 52. These different clinical balance scales focus on different aspects of balance. Based on a conceptual framework for functional balance tests in children[24], different aspects of balance can be defined based on the task constraints that relate to maintaining, achieving (i.e. object interaction) and restoring (i.e. obstacle negotiation) balance. The task constraints comprise three conditions: stability (i.e. in a certain position), quasi-mobility (i.e. during movement between postures), and mobility (i.e. during gait). Based on these definitions, it is suggested that PBS focuses on stability and quasi-mobility while maintaining and achieving balance, whereas BOTbal focuses on stability and mobility and BOTrsa only on mobility while maintaining balance. As such, different results may be found on the different balance scales after exergame-training as they focus on different aspects of balance.

Participants in the intervention group at VUMC performed two additional clinical balance scales; (iv) the Trunk Control Measurement Scale (TCMS) is a valid and reliable tool to assess static and dynamic sitting balance and is specifically developed for CP[25,26]. The subscale static sitting balance consists of five items (e.g., upright sitting and holding for 10 seconds). The subscale dynamic sitting balance consists of selective movement control and dynamic reaching, consisting of ten items (e.g., rotating the upper trunk with the head fixated in the starting position, reaching forward and sideways). As such, TCMS focuses on stability while maintaining and achieving balance based on clinical framework defined by Verbecque et al. (2014)[24]. Every task is graded on a 2-4 point ordinal scale, with a maximum score of TCMS is 58. (v) the Challenge-module was developed supplementary to the Gross Motor Function Measure to address its ceiling effect in high-functioning CP[27]. The Challenge-module evaluates 23 advanced gross-motor tasks, focusing on balance, speed and coordination, such as: walking on a balance beam, throwing and catching of a basketball, waving through pylons. Challenge-module was, thus, mostly focused on mobility while maintaining balance. Items were scored on a 5-point ordinal scale. Scores on all individual items (best of three trials) were summed to obtain the Challenge-module score, with a maximum score of 115.

Higher scores on each of the clinical scales indicated better balance control.

Statistical analysis

Power analysis of the expected changes (pre-post) in PBS indicated a required sample size of 18 children in the intervention group to detect a meaningful clinical difference of 3.66 points (SD of 3.7 [average of 2.7 and 4.7 for GMFCS level-I and II children] in CP[28], and assuming power of 80%, $\alpha=0.05$, and dropout of 10%).

Descriptive characteristics of the intervention and control group were compared using independent t-tests for continuous variables and chi-square tests for proportions.

To assess the differences on the raw scores of the balance scales between the intervention and control group before and after training (hypothesis 1), a mixed model ANOVA was performed with two factors including their interaction; one within-factor which assessed the effect of time (TIME, i.e. the pre-intervention versus the post-intervention effect), and one between-factor to assess the difference between the intervention and control group (INTERVENTIONGROUP). The interaction-effect (TIME*INTERVENTIONGROUP) provides information on whether changes over time (pre-post intervention) are different between the intervention and control group.

To assess the effect of 'distribution of CP-symptoms' on the effect of the intervention, a repeated measures ANCOVA was performed with two factors including their interaction; one within-factor which assessed the effect of time (TIME, i.e. the pre-intervention versus the post-intervention effect), and one covariate to assess the influence of the topographical distribution of CP-symptoms (CPDISTRIBUTION). The interaction-effect (TIME*CPDISTRIBUTION) provides information on whether the distribution of CP-symptoms influenced the results after the intervention (hypothesis 2).

To assess whether children with low baseline balance performance will improve more than children with high baseline balance performance (hypothesis 3) we perform three statistical steps. First we assessed the effect of 'baseline balance performance' on the effect of the intervention, using a repeated measures ANCOVA with two factors including their interaction; one within-factor which assessed the effect of time (TIME, i.e. the pre-intervention versus the post-intervention effect), and one covariate to assess

the influence of the baseline balance performance level (BASELINEBALANCE). The interaction-effect (TIME* BASELINEBALANCE) provides information on whether the baseline balance performance level influenced the results after the intervention (hypothesis 3, step 1). Next, when TIME*BASELINEBALANCE showed a significant effect, two groups were categorized based on baseline balance performance (high versus low baseline balance groups) using K-means cluster analysis with a convergence criterion of 0 and 2 clusters (hypothesis 3, step 2). Finally, to assess differences in the effect of training between participants with high versus a low baseline balance, a repeated measures ANOVA was performed with two factors including their interaction; one within-factor which assessed the effect of time (TIME, i.e. the pre-intervention versus the post-intervention effect), and one between-factor to assess the difference between the groups of participants (i.e. 1. participants in intervention group with a high baseline balance, 2. participants in intervention group with a low baseline balance level, and 3. the control group) (BASELINEGROUP). The interaction-effect (TIME* BASELINEGROUP) provides information on whether the control group, the high baseline balance group and the low baseline balance group respond different to the intervention (hypothesis 3, step 3).

Post-hoc pairwise comparisons were systematically applied in case of a significant interaction effect.

To assess the effect of training on the Challenge-module and TCMS in the subgroup that performed these assessments (at VUMC), dependent t-tests were used.

Statistics were performed with SPSS statistics 25 (IBM, USA). α was set at 0.05. Description of the results include statistical significance (p-value) and effect size (partial eta squared: η_p^2).

Results

Baseline demographics

Thirty-five children with CP were enrolled; 19 at UHG and 16 at VUMC. Recruitment ended in November 2017 at UHG, and June 2019 at VUMC. Patient characteristics were similar between the intervention and control group (table 1). See figure 1 for flow diagram of enrollment.

Table 1 near here

Figure 1 near here

Differences in balance between the intervention and control group and the effect of training (hypothesis 1)

No differences in balance scales outcome between the intervention group and control group were found (Figure 2; no main effect of INTERVENTIONGROUP; PBS:$p=0.255,\eta_p^2=0.039$; BOTbal:$p=0.059,\eta_p^2=0.104$; BOTrsa:$p=0.316,\eta_p^2=0.030$).

One balance scale (BOTrsa) significantly improved after the training period, but to a similar extent in both the intervention group and the control group (Figure 2; main effect of TIME for BOTrsa; $p=0.001,\eta_p^2=0.288$, but not for PBS or BOTbal - PBS:$p=0.248,\eta_p^2=0.040$; BOTbal:$p=0.200,\eta_p^2=0.049$).

There was no difference in balance scales outcome between the intervention group and control group at baseline and after 8 weeks (Figure 2; No interaction effect of TIME*INTERVENTIONGROUP; PBS:$p=0.248,\eta_p^2=0.040$; BOTbal:$p=0.374,\eta_p^2=0.024$; BOTrsa:$p=0.841,\eta_p^2=0.001$).

See Figure 2 for mean and standard deviation of the different outcome variables for the two groups.

Figure 2 near here.

Effect of training on Challenge-module and TCMS in VUMC subgroup (hypothesis 1)

Children with CP at the VUMC location showed a significant improvement in Challenge-module score after the intervention (Figure 2; p=0.003,η_p^2=0.502), but not for TCMS (Figure 2; p=0.150,η_p^2=0.153). It is, however, not clear what the natural evolution is for these balance outcomes as these were not performed in a control group.

See figure 2 for mean and standard deviation of the outcome variables.

Effect of topographical distribution of CP-symptoms on training (hypothesis 2)

Children with unilateral CP showed, on average, higher balance scores (i.e. better balance) compared to children with bilateral CP (mean ± standard deviation; PBS:55.4±0.8 vs 52.1±3.8; BOTbal:26.3±11.0 vs 16.7±7.2; BOTrsa:33.9±5.0 vs 17.3±11.4) as shown by the significant main effect of CPDISTRIBUTION for BOTbal and BOTrsa (BOTbal:p=0.027,η_p^2=0.204; BOTrsa:p=0.005,η_p^2=0.310). The result for PBS did not reach significance (PBS:p=0.066,η_p^2=0.145).

There was no difference in balance scales outcome between children with

unilateral CP and children with bilateral CP at baseline and after 6-8 weeks (No interaction effect of TIME*CPDISTRIBUTION; PBS:$p=0.373$,$\eta_p^2=0.036$; BOTbal:$p=0.127$,$\eta_p^2=0.103$; BOTrsa:$p=0.474$,$\eta_p^2=0.024$).

Effect of baseline balance performance on training (hypothesis 3, step 1)

Baseline balance level appeared to have a significant effect on the extent to which participants improved balance after training (based on the significant interaction effect of TIME*BASELINEBALANCE(covariate)). Baseline PBS score had a significant effect on PBS score and BOTbal score after training (PBS:$p<0.001$,$\eta_p^2=0.441$; BOTbal:$p<0.001$,$\eta_p^2=0.561$). Baseline BOTbal score had a significant effect on BOTbal score after training (BOTbal:$p<0.001$,$\eta_p^2=0.505$). Baseline BOTrsa score had a significant effect on BOTbal score after training (BOTbal:$p=0.005$,$\eta_p^2=0.308$).

As baseline PBS score showed the highest effect (lowest p-value and highest effect size) on balance after training, further analyses focused on this variable to subdivide the intervention group into a high and low balance baseline level group.

High and low baseline group determination using cluster analysis (hypothesis 3, step 2)

Based on the significant effect of baseline balance performance, a high baseline and low baseline balance performance group was determined for PBS in the intervention group. K-means cluster analysis revealed a high (PBS$\geq$50,n=19) and low (PBS<50,n=5) baseline balance performance group ($p<0.001$).

Comparison of effect of training between the high baseline balance group, the
low baseline balance group and the control group (hypothesis 3, step 3)

When the intervention group was divided into two groups based on PBS, we found that

children with a low baseline balance level significantly improved balance (PBS and

BOTbal) after exergame intervention, whereas children with a high baseline balance

level and the control group did not (Figure 3 & Table 2; significant interaction effect of

TIME*BASELINEGROUP; PBS:$p=0.003$,$\eta_p^2=0.304$; BOTbal:$p=0.008$,$\eta_p^2=0.258$).

Children with a low baseline (PBS) balance level did not improve significantly different

than children with a high baseline (PBS) and the control group balance level for BOTrsa

($p=0.379$,$\eta_p^2=0.059$). Children with a low baseline (PBS) balance level did not improve

significantly different than children with a high baseline (PBS) for TCMS ($p=0.800$

9,$\eta_p^2=0.006$), and Challenge-module ($p=0.354$,$\eta_p^2=0.072$).

See figure 3 and table 2 for mean and standard deviation of the PBS, BOTbal

and BOTrsa for the high and low baseline groups versus controls.

Figure 3 near here.

Table 2 near here.

Discussion

Based on the current results we cannot confirm our first hypothesis that children with

CP who received this exergame balance-training improved balance performance more

than children with CP who only received usual care. We could also not confirm our second hypothesis, as the effectiveness of this exergame-training to improve balance was not affected by topographical distribution of CP (i.e. unilateral or bilateral). However, we could confirm our third hypothesis that with this exergame-training, children from our current population with low baseline balance performance show more improvements than children with high baseline balance performance and than children who only received usual care.

Regarding hypothesis one ('children with CP who receive exergame-training will improve balance performance more than those who do not'), on average (for the total group of CP), we did not find an effect of this exergame-training on changes in scores on balance scales compared to a control group. We did find a significant increase in the Challenge-module from pre- to post-training in a subgroup of participants who performed this balance scale (at VUMC); i.e. difference in mean of 3.42 points. This change, however, did not reach the minimal detectable change of 4.47 points[29]. Additionally, for this outcome, the natural evolution over 8 weeks is unknown as the control group did not perform this test. Moreover, significant changes after training were also found for BOTrsa, but the improvements were similar in the intervention and control group. For this reason, this improvement does not seem to be related to the exergame-training, indicating that this change can be due to natural development or the received usual care program of the included participants. Additionally, it may be the case that this outcome may be more related to fitness as the original scale was developed as a measure of fitness rather than balance. These negative results for the different balance outcomes are in accordance with previous studies that did not find an effect of exergame-training[15–17]. However, these negative results could be caused by

subgroup or interindividual variability in change of balance performance over time. Based on the individual data (Figure 2), it is clear that there is a high interindividual variability regarding the effectiveness of the exergame-intervention; some individuals clearly show an increase in balance whereas others show no improvement or even a decrease. Subgroup analyses can provide insights in whether the individuals who show an improvement after training possess common features.

Regarding hypothesis two ('children with bilateral CP who receive exergame-training will improve balance performance more than children with unilateral CP'), the current results showed that the topographical distribution of CP-symptoms (i.e. unilateral or bilateral CP) did not influence the effectiveness of this type of balance training, despite the differences in balance between unilateral and bilateral children with CP. This is an unexpected finding as children with unilateral CP have an unaffected side which could compensate for the affected side while in children with bilateral CP both the lower limbs are affected. This subgroup analysis indicated that distribution of the CP-symptoms did not influence the effectiveness of the exergame-training.

On the other hand, regarding hypothesis three ('children with low baseline balance performance will improve more than children with high baseline balance performance'), we did find that children from the current population with poorer balance performance at baseline did improve their balance after this exergame-training focused on balance (irrespective of the distribution of CP-symptoms). As such, classification/grouping based on baseline balance level differentiated our population in a different manner than grouping based on topographical distribution of CP-symptoms; as both children with bilateral and unilateral CP may present with a high or low baseline

balance level, topographical distribution of CP-symptoms does not appear to reflect functional balance performance. Children with a baseline score of less than 50 points on the PBS, improved on average 4 points on the PBS and 5 points on the BOTbal scale after exergame-training. This change is significant as the minimal detectable change is exceeded for both PBS (1.59 points) and BOTbal (1.14 points)[30,31]. Specifically, 60% (3/5) children in the intervention group with a low baseline level increased 2 points or more on the PBS and the BOTbal. None of these children showed a decrease of 2 or more points. In the high baseline balance group, however, only 15.8% (3/19) children increased 2 or more points on the PBS and 31.6% (6/19) children increased 2 or more points on the BOTbal. While 15.8% (3/19) and 57.9% (11/19) of them showed a decrease of 2 or more points on the PBS and BOTbal, respectively. In the control group, 9.09% (1/11) and 45.5% (5/11) children showed an increase of 2 points or more on the PBS and the BOTbal, respectively. Whereas, 18/2% (2/11) children in the control group showed a decrease of 2 or more points on the BOTbal. Furthermore, the change seen in this low baseline CP group after exergame-training is clinically meaningful as the minimal clinical important difference for PBS (3.66 points) and BOTbal (0.57 points) were exceeded[30,31]. The finding that children with CP with a low performance at baseline may benefit more from training than children with a high performance at baseline, is in line with results from upper limb training in CP[20,21], and sitting balance training using exergames in children with CP who were hospitalized after lower limb surgery[14].

Our results suggest that promising or negative effects in previous studies may be the resultant of a selection bias of included participants with a low or high baseline balance performance, respectively, as most previous studies on the effectiveness of

exergame balance training had low sample sizes. To the best of our knowledge, the current study has included the largest sample of children included in an exergame-intervention focusing on balance improvement, addressing the need for more studies with higher sample sizes as indicated in previous reviews[6,18,19]. However, next to a low sample size, other factors may influence the effectiveness of exergame-training as well. Possible future directions for research to investigate different factors are suggested. To date it is not yet known to what extent different types of exergaming affect training outcome. Furthermore, patient motivation is crucial for compliance especially for longer training periods. As such, exergames with an aspect of fun included may result in increased motivation and compliance[7]. Given the interindividual differences in effect of training, it appears that patient-specific adjustments of the (exergame)training are required to (1) train on specific aspects of balance control (e.g. reactive standing balance, anticipatory movements), and (2) present a fun and motivating environment of specific interest to the patient. Motivation does not only affect the training but also the assessment of balance in children with CP. It is not expected that balance performance shows a natural evolution over a 6-8 week period. However, the individual results of the participants in the control group (Figure 2) do show variations when assessed the second time (after 8 weeks) both in the positive and negative direction (for PBS and BOTbal). This suggests that it is not a learning effect of the test for PBS and BOTbal, as then the second test would result in a better performance (as is the case for BOTrsa). Therefore, it seems the child's motivation to perform the tests may have affected the results of the PBS and BOTbal. Of course, this may not only be the case for the control group but also for participants in the exergame-intervention group. Alternatively, a change in the usual care of the patient over the course of the study could have influenced the results as well. However, we believe this

will not have been an important issue as participants and their parent were asked not to change the usual care and if this would happen that they would contact the research team (none have indicated such a change). Irrespective of the possible reason, individual variations were present. The individual differences between baseline and after 6-8 weeks for PBS in the control group ranged from -1 to +2, but did not reach the minimal clinical important difference for PBS. For BOTbal, the individual variation of the control group was much higher, ranging from -5 to +8, and did reach the minimal clinical important difference for BOTbal. After intervention in the exergame-intervention group, 5 children showed a negative effect on the PBS (up to -2), which was not clinically important. However, for BOTbal, half of the exergame-intervention group presented with a negative effect after training (up to -13), indicating that the BOTbal outcome is too variable to use as balance performance outcome in children with CP.

The current results indicate that baseline PBS level was a better factor to distinguish groups for possible benefit of this exergame balance training than BOTbal and BOTrsa. Children with a low baseline PBS level improved at both PBS and BOTbal after this training, whereas baseline BOTbal level or BOTrsa level only influenced BOTbal after exergame-training. As indicated in the methods section, PBS focuses on stability and quasi-mobility while maintaining and achieving balance, whereas BOTbal focuses on stability and mobility and BOTrsa on mobility while maintaining balance. When assessing the type of balance (i.e. task constraints[24]) trained during exergaming, it appears that mainly quasi-mobility is trained while maintaining and achieving balance, as the tennis, bowling and soccer games required performing a transfer of body position (e.g. a lunge and returning, or a step while swinging the arm).

As such, PBS appears to better reflect the type of balance trained with the exergames. Based on the results of the clinical balance scales of the low baseline balance group, it appears that this exergame-training focused on balance results in task-specific training with limited to no transfer to other types (or task constraints) of balance, such as mobility (e.g. balance during walking). Similarly, we also did not find a transfer of the exergame-training with a focus on standing balance to improvements on the TCMS which assesses sitting balance.

Study limitations

When interpreting the results of the current study, one should take into account some limitations. We aimed to randomize the allocation of the participants to the control group of intervention group at both locations. We were, however, required to start with the recruitment of participants before the X-box consoles were delivered (to reduce project delays). For that reason, participants who started at the beginning of the project were allocated to the control group. Additionally, due to slow recruitment of patients, and a resulting project delay, we opted towards the end of the project to mainly allocate participants to the intervention group. These allocation procedures were not related to clinical characteristics of the patients. Blinding of participants to group allocation was not an option, as persons in the intervention group received X-box consoles, while persons in the control group did not. Additionally, the capacity of the research group did not allow blinding of the researchers performing balance assessments. Nevertheless, the balance scales used required objective scoring. When a significant effect of baseline balance score as covariate was found, the intervention group was divided in a high and low baseline balance group based on K-means cluster analysis. Future research in a larger sample of children with CP that receive exergame balance training, should focus on the validity of this threshold or other methods to determine a threshold of when

exergame training is meaningful. Combining the results of two centers (VUMC and UHG) allowed an increase in sample size, but also increased variability. At UHG children were asked to train for 8 weeks, while at VUMC this was 6 weeks. On average, however, there was no significant difference in the hours played between the groups, suggesting that children at UHG either trained less than 30min per day or did not train 5 days per week. Nevertheless, as the hours played was similar between the groups, both groups have trained for a similar amount of time (table 1). Furthermore, different investigators assessed the balance scales in the two institutions. This was believed not to be an important issue as the interrater reliability of the PBS and BOTbal are very high(ICC=0.98)[32,33]. A weakness of the current study is that for hypothesis 2 and 3 the large initial sample was divided into (unbalanced) smaller subgroups. These analyses provide important insights into possible reasons of the negative findings on the group level which should be investigated in future large RCTs, and should be interpreted with caution. Some children may have had more experience with gaming using the Kinect Xbox One console, or a different console, than others. We did not assess this aspect, and therefore, could not investigate the possible influence of previous gaming experience on the effectiveness of this exergame-training. Another aspect that deserves further examination, but was not part of the current study, is the effect of comorbidities on exergame balance-training and balance performance; for instance, how do ADHD, autism, IQ, and visual or auditory problems affect this type of intervention. Importantly, readers should not generalize the current findings to other types of exergame-training nor to a different (albeit similar) population. These findings are specific for home-based videogame training using the Kinect Xbox One game Kinect sports rivals (and specifically for three sub-games: tennis, soccer and bowling), as different games may have a focus on different aspects of balance or different type of

visual feedback which may influence effectiveness. Additionally, these findings are specific to the currently investigated population (children with bilateral and unilateral spastic CP with GMFCS level I-II).

Conclusions

This home-based exergame balance training using Kinect X-box games did not improve functional balance for all included children with spastic unilateral or bilateral CP with a GMFCS level-I or II. However, children with a low baseline balance performance (PBS-score of 50 or less), did show significant and clinically important improvements on the PBS and BOTbal after this exergame balance-training for 6-8 weeks. This exergame balance-training did not result in a transfer to improvements in sitting balance or balance during mobility for children with a low baseline function. The topographical distribution of CP-symptoms (i.e. unilateral or bilateral CP) did not affect the effectiveness of the training.

References

1. Odding E, Roebroeck ME, Stam HJ. The epidemiology of cerebral palsy: incidence, impairments and risk factors. Disabil Rehabil. 2006 March 28;28:183–91.

2. Aisen ML, Kerkovich D, Mast J, Mulroy S, Wren T Al, Kay RM, Rethlefsen SA. Cerebral palsy: clinical care and neurological rehabilitation. Lancet Neurol. 2011;10:844–852.

3. Carlberg EB, Hadders-Algra M. Postural dysfunction in children with cerebral palsy: Some implications for therapeutic guidance. Neural Plast. 2005;12:221–228.

4. Bruijn SM, Millard M, van Gestel L, Meyns P, Jonkers I, Desloovere K. Gait

stability in children with Cerebral Palsy. Res Dev Disabil. 2013;34.

5.   Franki I, Desloovere K, De Cat J, Feys H, Molenaers G, Calders P, Vanderstraeten G, Himpens E, Van Broeck C. The evidence-base for basic physical therapy techniques targeting lower limb function in children with cerebral palsy: a systematic review using the International Classification of Functioning, Disability and Health as a conceptual framework. J Rehabil Med. 2012 May;44:385–95.

6.   Chen Y, Fanchiang HCD, Howard A. Effectiveness of virtual reality in children with cerebral palsy: A systematic review and meta-analysis of randomized controlled trials. Phys Ther. 2018 January 1;98:63–77.

7.   Meyns P, Roman de Mettelinge T, van der Spank J, Coussens M, Van Waelvelde H. Motivation in pediatric motor rehabilitation: A systematic search of the literature using the self-determination theory as a conceptual framework. Dev Neurorehabil. 2017.

8.   Booth AT, Buizer AI, Harlaar J, Steenbrink F, van der Krogt MM. Immediate Effects of Immersive Biofeedback on Gait in Children With Cerebral Palsy. Arch Phys Med Rehabil. 2019 April 1;100:598–605.

9.   Holden MK. Virtual environments for motor rehabilitation: review. Cyberpsychol Behav. 2005;8:187–211; discussion 212-219.

10.   Jung SH, Song SH, Kim SD, Lee K, Lee GC. Does virtual reality training using the Xbox Kinect have a positive effect on physical functioning in children with spastic cerebral palsy? A case series. J Pediatr Rehabil Med. 2018;11:95–101.

11.   Gatica-Rojas V, Méndez-Rebolledo G, Guzman-Muñoz E, Soto-Poblete A, Cartes-Velásquez R, Elgueta-Cancino E, Lizama LEC. Does Nintendo Wii Balance Board improve standing balance? A randomized controlled trial in

children with cerebral palsy. Eur J Phys Rehabil Med. 2017 August 1;53:535–544.

12. Tarakci D, Ersoz Huseyinsinoglu B, Tarakci E, Razak Ozdincler A. Effects of Nintendo Wii-Fit® video games on balance in children with mild cerebral palsy. Pediatr Int. 2016;58:1042–1050.

13. Atasavun Uysal S, Baltaci G. Effects of Nintendo Wii ™ Training on Occupational Performance, Balance, and Daily Living Activities in Children with Spastic Hemiplegic Cerebral Palsy: A Single-Blind and Randomized Trial. Games Health J. 2016 October 1;5:311–317.

14. Meyns P, Pans L, Plasmans K, Heyrman L, Desloovere K, Molenaers G. The Effect of Additional Virtual Reality Training on Balance in Children with Cerebral Palsy after Lower Limb Surgery: A Feasibility Study. Games Health J. 2017;6.

15. Ramstrand N, Lygnegård F. Can balance in children with cerebral palsy improve through use of an activity promoting computer game? Technol Heal Care. 2012;20:501–510.

16. Jelsma J, Pronk M, Ferguson G, Jelsma-Smit D. The effect of the Nintendo Wii Fit on balance control and gross motor function of children with spastic hemiplegic cerebral palsy. Dev Neurorehabil. 2013;16:27–37.

17. Sajan JE, John JA, Grace P, Sabu SS, Tharion G. Wii-based interactive video games as a supplement to conventional therapy for rehabilitation of children with cerebral palsy: A pilot, randomized controlled trial. Dev Neurorehabil. 2017 August 18;20:361–367.

18. Wu J, Loprinzi PD, Ren Z. The rehabilitative effects of virtual reality games on balance performance among children with cerebral palsy: A meta-analysis of

randomized controlled trials. Int J Environ Res Public Health. 2019 November 1;16.

19. Warnier N, Lambregts S, Port I Van De. Effect of Virtual Reality Therapy on Balance and Walking in Children with Cerebral Palsy: A Systematic Review. Dev Neurorehabil. 2019.

20. Klingels K, Feys H, Molenaers G, Verbeke G, Van Daele S, Hoskens J, Desloovere K, De Cock P. Randomized trial of modified constraint-induced movement therapy with and without an intensive therapy program in children with unilateral cerebral palsy. Neurorehabil Neural Repair. 2013;27:799–807.

21. Sakzewski L, Ziviani J, Boyd RN. Best responders after intensive upper-limb training for children with unilateral cerebral palsy. Arch Phys Med Rehabil. 2011 April;92:578–584.

22. Yi SH, Hwang JH, Kim SJ, Kwon JY. Validity of pediatric balance scales in children with spastic cerebral palsy. Neuropediatrics. 2012;43:307–313.

23. Deitz JC, Kartin D, Kopp K. Review of the Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2). Phys Occup Ther Pediatr. 2007;27:87–102.

24. Verbecque E, Lobo Da Costa PH, Vereeck L, Hallemans A. Psychometric properties of functional balance tests in children: a literature review. Dev Med Child Neurol. 2015;57:521–529.

25. Pham HP, Eidem A, Hansen G, Nyquist A, Vik T, Saether R. Validity and Responsiveness of the Trunk Impairment Scale and Trunk Control Measurement Scale in Young Individuals with Cerebral Palsy. Phys Occup Ther Pediatr. 2016 November;36:440–452.

26. Marsico P, Mitteregger E, Balzer J, van Hedel HJA. The Trunk Control

Measurement Scale: reliability and discriminative validity in children and young people with neuromotor disorders. Dev Med Child Neurol. 2017;59:706–712.

27. Wilson A, Kavanaugh A, Moher R, McInroy M, Gupta N, Salbach NM, Wright FV. Development and pilot testing of the challenge module: A proposed adjunct to the gross motor function measure for high-functioning children with cerebral palsy. Phys Occup Ther Pediatr. 2011;31:135–149.

28. Duarte ND a C, Grecco LAC, Franco RC, Zanon N, Oliveira CS. Correlation between Pediatric Balance Scale and Functional Test in Children with Cerebral Palsy. J Phys Ther Sci. 2014;26:849–53.

29. Wright FV, Lam CY, Mistry B, Walker J. Evaluation of the Reliability of the Challenge when used to Measure Advanced Motor Skills of Children with Cerebral Palsy. Phys Occup Ther Pediatr. 2018;38:382–394.

30. Chen C, Shen I, Chen C, Wu C, Liu W, Chung C. Validity, responsiveness, minimal detectable change, and minimal clinically important change of Pediatric Balance Scale in children with cerebral palsy. Res Dev Disabil. 2013 March;34:916–22.

31. Wuang YP, Su CY. Reliability and responsiveness of the Bruininks-Oseretsky Test of Motor Proficiency-Second Edition in children with intellectual disability. Res Dev Disabil. 2009;30:847–55.

32. Franjoine MR, Gunther JS, Taylor MJ. Pediatric balance scale: A modified version of the Berg Balance Scale for the school-age child with mild to moderate motor impairment. Pediatr Phys Ther. 2003;15:114–128.

33. Bruininks R, Bruininks B eds. Bruininks-Oseretsky Test of Motor Proficiency. 2nd ed. Minneapolis: NCS Pearson; 2005.

Table 1. Characteristics of participants in the intervention and control group.

	Intervention group	Control group	Chi^2 or t-statistic	p-value
N	24	11		
Institution (VUMC/UHG)	14/10	2/9		
Gender (M/F)	19/5	9/2	Chi^2: 0.033	0.856
GMFCS level (I/II)	9/15	3/8	Chi^2: 0.350	0.554
Distribution (uni-/bilateral)	5/19	2/9	Chi^2: 0.033	0.856
Age (years)	10.95 (2.34)	9.78 (1.06)	t: 1.578	0.124
Weight (kg)	38.07 (13.83)	32.92 (5.93)	t: 1.180	0.247
Height (m)	1.46 (0.15)	1.38 (0.09)	t: 1.548	0.131
Hours played (hours)	20.9 (6.9)°	0.0 (0.0)	**t: 9.933**	**<0.001**

Continuous variables are presented as follows: mean (standard deviation). VUMC = Amsterdam UMC location VUmc, UHG = University Hospital Ghent, N = number of participants, M/F = male/female, GMFCS = Gross Motor Function Classification System. Chi-square tests were performed for Gender and GMFCS. Independent t-tests were performed for Age, Weight, Height and Hours played. °: separate mean (standard deviation) for VUMC: 21.0 (4.7) and UHG: 20.7 (9.2).

Table 2. Mean and standard deviations of the PBS, BOTbal and BOTrsa for the high baseline balance group, low baseline balance group and the control group, including post-hoc pairwise comparisons.

	High baseline balance group	Low baseline balance group	Control group	Post-hoc high vs low p-value	Post-hoc low vs control p-value	Post-hoc high vs control p-value
	N=19	*N=5*	*N=11*			
PBS						
Baseline	54.11 (2.00)	45.20 (3.90)	54.27 (3.47)	**<0.001**	**<0.001**	0.876
Post-intervention (after 6-8 weeks)	54.37 (2.57)	49.00 (2.12)	54.27 (3.64)	**0.001**	**0.002**	0.931
Post-hoc pre vs post p-value	0.574	**<0.001**	1.000			
BOTbal						
Baseline	21.32 (9.62)	8.00 (4.90)	23.91 (8.43)	**0.005**	**0.002**	0.442
Post-intervention (after 6-8 weeks)	20.32 (8.03)	13.40 (2.30)	25.73 (7.94)	0.077	**0.005**	0.067
Post-hoc pre vs post p-value	0.284	**0.005**	0.143			
BOTrsa						
Baseline	23.32 (11.13)	7.00 (4.18)	24.27 (10.64)	-	-	-
Post-intervention (after 6-8 weeks)	24.63 (12.33)	10.20 (7.36)	26.18 (11.87)	-	-	-
Post-hoc pre vs post p-value	-	-	-			

Variables are presented as follows: mean (standard deviation). PBS = Pediatric Balance Scale, BOTbal = balance subscale of the Bruininks-Oseretsky Test of Motor Proficiency-2, BOTrsa = running speed & agility subscale of the Bruininks-Oseretsky Test of Motor Proficiency-2.
Note that the post-hoc tests were not performed for BOTrsa as the interaction effect did reach statistical significance.
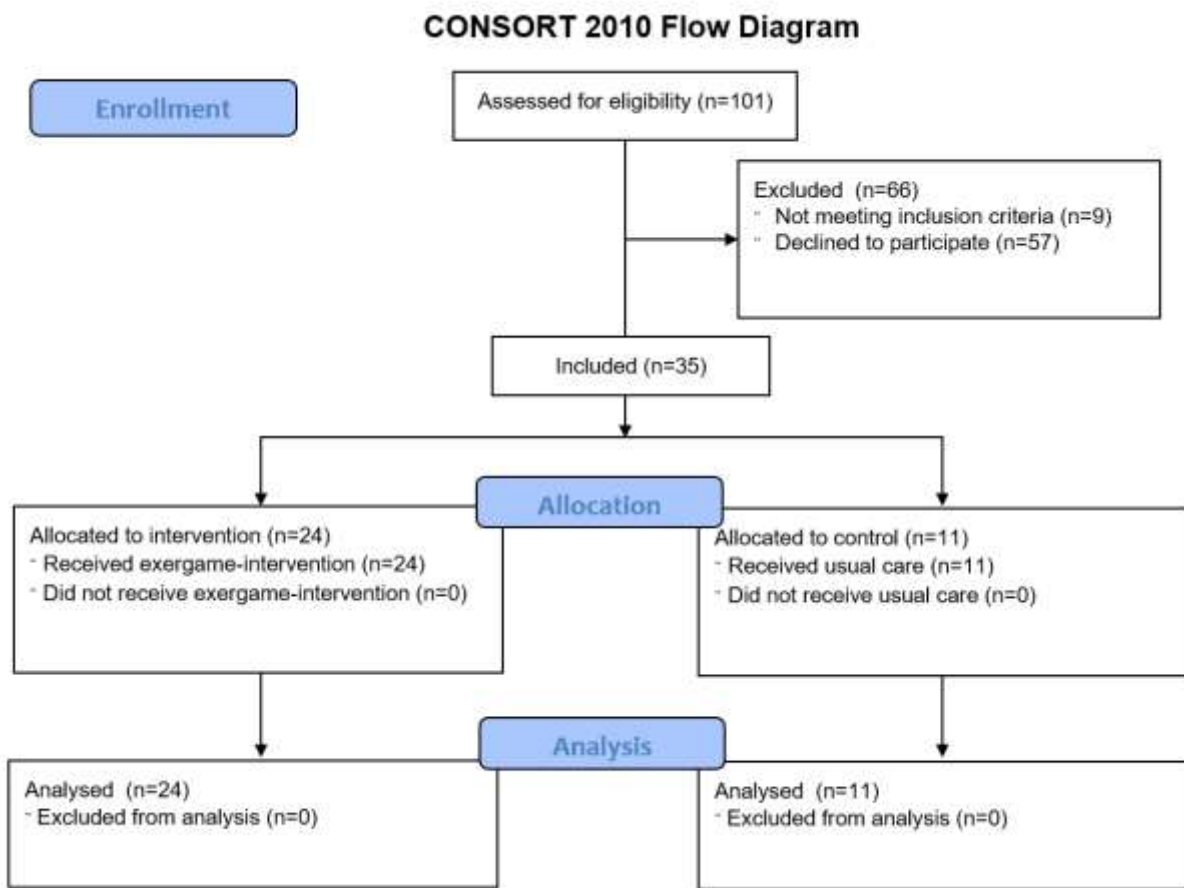
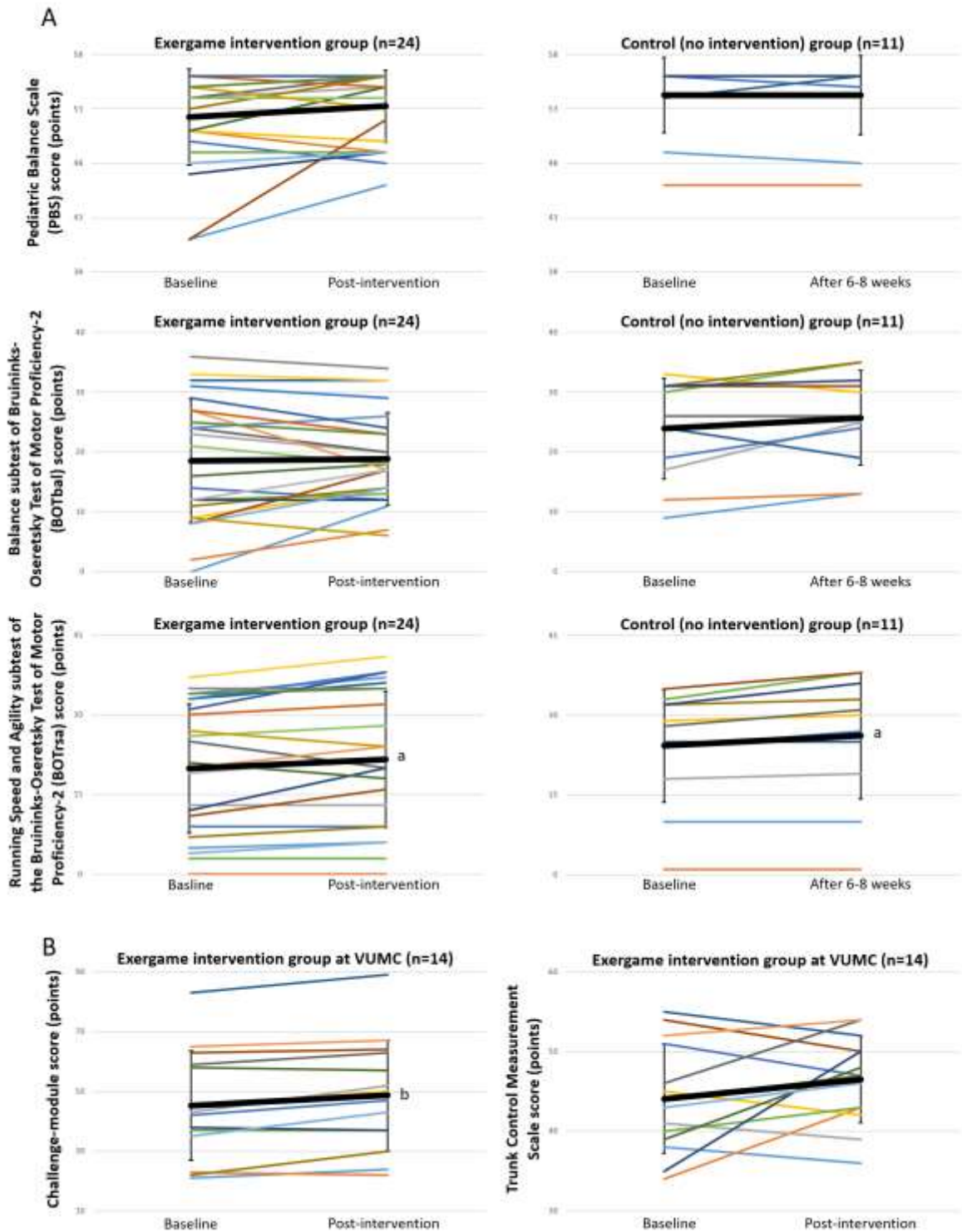Figure 1. Consort flow diagram addressing the enrolment of participants in the study.

Figure 2. Results of the exergame intervention in the intervention group with respect to the natural evolution of balance in the control group are provided in pane A (for PBS, BOTbal and BOTrsa). No effect of the exergame intervention was found, as both the exergame and control group did not show differences between baseline and after 6-8 weeks for PBS and

BOTbal. For BOTrsa, a difference between baseline and after training was found in the intervention group, but a similar effect was found for the control group between baseline and after 6-8 weeks (see 'a' in figure 2). For a subgroup of participants (at VUMC), the effect of exergame intervention was assessed on the Challenge-module and TCMS scales. Results of the exergame intervention in the intervention group in this subgroup is provided in pane B. Results on the Challenge-module show that the VUMC subgroup in the exergame intervention group significantly scored better after training than at baseline (see 'b' in figure 2).

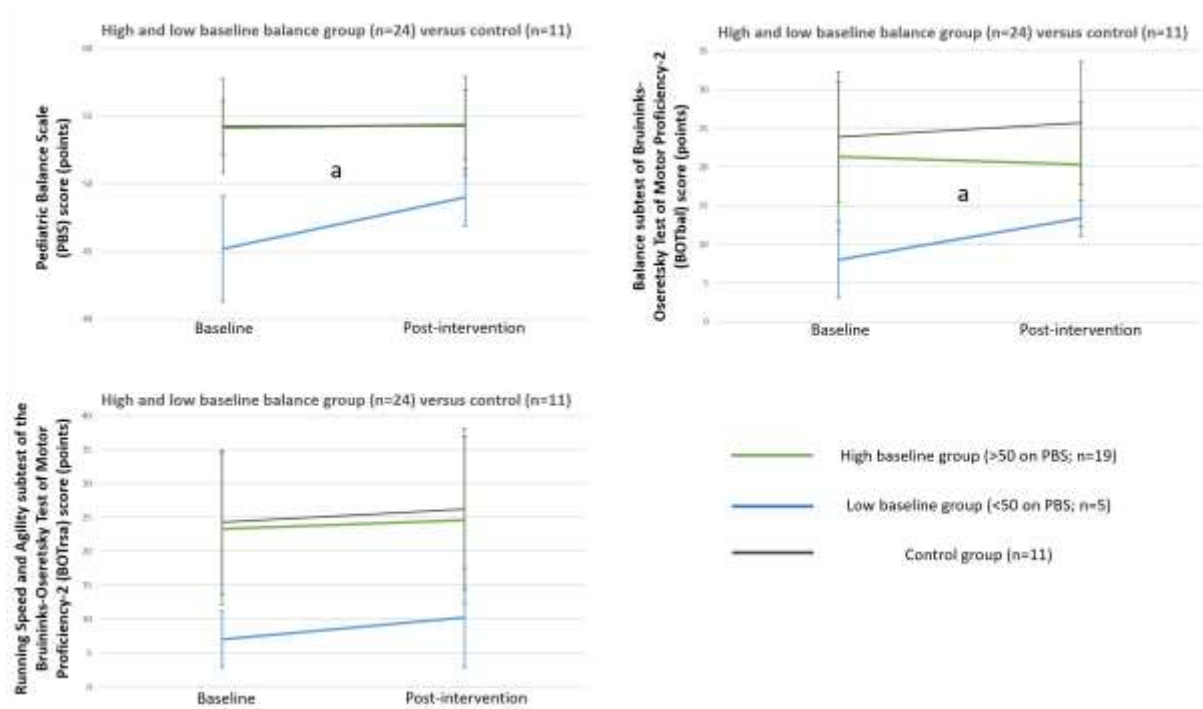Colored lines represent individual data. Thick black lines represent the mean with standard deviations.

Figure 3. Results at baseline and post-intervention on the PBS, BOTbal and BOTrsa scales for the high (>50 points) and low baseline group (<50 points) of the intervention group (based on the score on the PBS). Green and blue lines represent the mean with standard deviations of the high and low baseline group, respectively. The grey line represents the mean with standard deviations of control group. The low balance baseline group showed a significant increase in balance outcome for PBS and BOTbal after intervention compared to the high balance baseline group and the control group (a). This was not the case for BOTrsa.