PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Can an online exercise prescription tool improve adherence to home exercise programmes in children with cerebral palsy and other neurodevelopmental disabilities? A randomised controlled trial
AUTHORS	Johnson, Rowan; Williams, Sian; Gucciardi, Daniel; Bear, Natasha; Gibson, Noula

VERSION 1 – REVIEW

REVIEWER	Nicholas Taylor La Trobe University, Australia I have recently been a co-author on another manuscript submitted for consideration for publication with author Natasha Bear.
REVIEW RETURNED	21-May-2020

GENERAL COMMENTS	The authors are congratulated on completing a randomised controlled trial in the community with young people with cerebral palsy. The results found no significant difference in adherence or other primary outcomes between participants receiving exercise instructions with the aid of an online tool compared with those receiving exercise instructions with paper-based methods. As the confidence intervals about the primary outcome of adherence were less than 34% (sample size estimation) the authors can therefore conclude that the use of the online prescription tool was not superior to traditional paper-based methods of prescribing exercises. One concern is the number of apparent minor variations from the protocol in the outcome (e.g. proportion of days exercise) and in the analysis (e.g. planned incidence rate ratios for adherence) – see below for more detail. Although overall the trial has been completed according to the protocol it is important that any minor variations be declared and explained. Another concern was the fact that the analysis of the primary outcomes appears to have been completed as a per protocol analysis rather than as intention to treat as planned. For example, if there were log-book data of any participants withdrawing during the program their data should be included in the analysis. Some specific comments 1. Abstract page 3, line 18: Stated that n= 54 were recruited, but numbers with CP and other disabilities only adds up to n=46. 2. Abstract page 3, line 17: Please specify the primary and secondary outcomes separately. 3. Abstract page 3, line 22: Clarify the sample size on which analysis was completed (n=46). It appears that a per protocol analysis has been reported, only including participants who completed the 8-week program.

	4. Strengths and limitations, page 4, line 15: Stated that recruitment target of n=58 was not attained (n=46). However, in the trial registry sample size estimation was n=66 including an allowance for attrition. Also, in the Abstract it stated that n=54 were recruited. It appears n=46 was the number completing the intervention. 5. Data analysis page 9, line 24: According to protocol, adherence also measured as number of days exercised. This outcome has not been reported. 6. Data analysis page 9, about line 26: There appear to be some minor variations from the planned analysis of adherence. For example, incidence rate ratios were planned but not reported. Also, it was planned to adjust for number of follow up visits between the groups but this has not been reported. Also, it appears the analysis of weekly adherence, appropriately using a mixed effects model, was not planned - is that right? 7. Results page 10, about line 10: Please report the characteristics of participants who did not complete the intervention. 8. Results section: Results on primary outcomes were reported on 44, 46 and 40 participants, respectively. As presented, this presents as a per-protocol analysis rather than an intention to treat analysis, only including data for those who completed the intervention. If available, data from participants who dropped out during the program should be included. 9. Results: How many follow-up visits were there by the physiotherapists in each group – this does not appear to be reported. 10. Results: What were the characteristics of typical exercise programs? E.g. How many exercises? Dose? Time to complete prescribed program. Were these factors similar between the groups? 11. Patient flow figure: Should include numbers analysed for exercise performance data. 12. Discussion page 17: Can you propose why the intervention did not lead to hypothesised improvements in adherence and other primary outcomes? Could it be that the mode of exercise instruction is less important than the therapist who prescribes the exercises – as alluded to in the introduction. What are the clinical implications of the findings?

REVIEWER	Robert Ware Griffith University, Australia
REVIEW RETURNED	22-Jun-2020

GENERAL COMMENTS	Thank you for the opportunity to review this interesting manuscript I have some minor comments which I hope will improve readability Abstract/Results – between group difference for the primary outcome should be reported in abstract (in this case between-group mean difference (95%CI) for adherence) Abstract/Results – does “There was a downward trend in adherence in both groups of -2.3% (95% CI: -3.3 to -1.3, p<0.001) for each consecutive week” refer to both groups individually or both groups combined? – was there a between-group difference in rate of adherence over time? It should be noted in the abstract that this was an exploratory analysis (i.e. it was not pre-specified)

	Abstract/Results (and throughout manuscript) - Usually p-values are reported to 2 decimal places to p=0.01, then 3 decimal places to P<0.001 Article summary – the article summary that follows the abstract would be more useful if it included the key results of the study Methods – Methods section should be ordered according to the CONSORT guidelines – in particular the sample size section (item 7 in CONSORT 2010) should appear between Outcomes (item 6) and Randomisation (item 8) Methods – I was expecting to see an ethics statement regarding institutional approval towards the start of the Methods section. Methods – ANZCTR trial registration number should be included in first paragraph of Methods Methods – how were children stratified according to the Functional Mobility Scale? What cut-off was used? Methods/Data analysis – note that “Stata” is the name of the software rather than an acronym so should not be all upper-case Results/Table 1 – the Functional Mobility Scale and GMFCS (which is not defined) are listed in Table 1. Could these two scales be (briefly) defined in the Methods please? What do they measure and how are they ordered (e.g. does scoring a “1” mean the child has very good or very poor function)? Results/Adherence – rather than reporting the between group difference as “P>0.05” in the text, could the mean difference and 95%CI be reported instead please Results/Table 2 – could mean difference (95%CI) be reported for the ‘Adherence: Self-report” questions Results/Adherence and goal attainment – could 95%Cis be reported for the estimates of the correlation coefficient please

REVIEWER	Georgina Clutterbuck Western Sydney University, Australia
REVIEW RETURNED	15-Jul-2020

GENERAL COMMENTS	Thank you for the opportunity to review your paper. The role of technology in exercise prescription and program adherence is an important area of research and your work has provided important groundwork for their application in the population of children with disabilities. Overall I found the paper to be clear, concise and well designed, especially with the published protocol. I hope that the following feedback is helpful in strengthening the paper for publication. Abstract: 1. L13: Method- in the first sentence it reads like the community therapy was completed using Physitrack or paper-based methods, rather than the home exercise program. Strengths and Limitations: 1. Concise and well communicated

	2. L17- It would be beneficial to name the tools here. Introduction: 1. L15: It would be beneficial here to state that this definitions is specifically for children with disabilities as it may differ from those devised for other population groups (i.e. parental participation would not be relevant for most adults). 2. L32: You have clearly identified how mHealth could be used to capitalise on the facilitators of adherence based on this research. It would also be beneficial to include some of the barriers to adherence that were identified in this research that mHealth could target, or which may be out of scope for Health. Methods: 1. P8, L44: This sentence has some issues with plurality etc- a possible suggested edit: BoNTA injections were not an exclusion criterion because participation in an exercise program following BoNTA is recommended practice. 2. P9, L45: Did PTs use the library videos or customise their videos- do you have access to details about the proportion that used each type? 3. Further details about the number of exercises and type of exercises (and if they were similar in both groups) would improve the report of the intervention. 4. The protocol details that participants would receive weekly reminders, as well as continued intervention by the treating physiotherapist at the clinic or home. It would be beneficial to detail the number of home and centre visits received and if there was any difference between groups. 5. P10, L14: When were the three time points of that the COEP recorded? 6. P10, L19: Extra comma after “we also measured” 7. P11, L24: Adherence data were (not was) 8. The journal has requested particular attention be paid to certain CONSORT items. One of these is allocation concealment- it is not clear if/how the number sequence was concealed after it was randomly generated so as to conceal the allocation of the next enrolled participant. Results: Clearly reported. Discussion: 1. The paper would benefit from further discussion around the reasons that no difference between conventional and Physitrack adherence was found. In addition to the discussion around the responsiveness of the COEP, could the completion of the COEP (i.e. introducing filming of the home program to those who otherwise would only have a paper program) have influenced the feeling of being observed in the control group, thereby confounding the results? This may add to the comment regarding return of logbooks acting as an exercise reminder. There may have also been barriers to home exercise program adherence reported in the literature which continue with both methods. 2. Considering that 5% of participants were identified as potentially causing harm by doing their exercises, it may be appropriate to suggest that physiotherapists could use video to identify these children and therefore decrease harmful home exercise practice. 3. While the results did not support the initial hypothesis that Physitrack would be superior, the outcome of being equally as effective as conventional paper-based programs indicates that

	therapists can confidently use these methods based on the individual preferences and circumstances of their patients. Thank you again for the opportunity to review this paper. I look forward to seeing it published.

REVIEWER	Ferdinand von Walden Karolinska Institutet Sweden
REVIEW RETURNED	27-Jul-2020

GENERAL COMMENTS	The authors have completed an ambitious project to investigate the adherence to and effectiveness of an 8-week home exercise program for children with disabilities delivered using Physitrack or paper-based methods. Although the study is technically correct and adheres to the CONSORT guidelines for presenting data and methods I have major concerns as the study is under-powered. In the method section, the authors report the results of a power calculation that states the retirement of n=58 participants to be able to detect the expected difference between the two methods of program delivery. To allow for a few drop outs/non-finishers the aim was to recruit n=66. The actual number of participants that completed the trial was n=46. Given that no difference was detected between the programmes, the fact that the study is underpowered represents a problem in my view. Looking at the data, some data points suggest a better adherence in the intervention group (e.g. log book findings) that might have proved significant in a properly powered study.The fact that the study is underpowered to answer both primary and secondary outcomes makes me hesitant to recommend this manuscript for publication. The fact that the study is underpowered is not properly discussed in the discussion section. Minor comments: I suggest that the authors should reconsider the title of the manuscript, as a large proportion of the participants were non-CP. Please make the number of study participants more clear in the abstract. It currently says fifty-four participants were recruited (n=32 CP and n=14 other).However the stated numbers (n=x) refers to the number of participants that completed the trial. The number of participants in the intervention group is not consistent (22, 24) Participants: What was the rational for mixing deurodevelopmental diagnosis with such variable motor component? It's stated in the methods section that "recruitment period was determined by funding timeline agreements". So the primary reason for stopping the study before it was sufficiently powered was financial? Was prescribed oral medication documented? How many of the kids with CP received BTX during the intervention? Were these participants evenly distributed between the groups?

	Discussion: The conclusion of the study is overly positive in my opinion. First, the study is underpowered to make a good interpretation of the data. Second, as is, the study indicates that the development of an app such as Physiotrack seems unnecessary as the same outcomes can be obtained by paper-based documentation. Based on the available data, the discussion and conclusion is somewhat unbalanced.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Nicholas Taylor
Institution and Country: La Trobe University, Australia
Please state any competing interests or state 'None declared': I have recently been a co-author on another manuscript submitted for consideration for publication with author Natasha Bear.

General comment:
The authors are congratulated on completing a randomised controlled trial in the community with young people with cerebral palsy. The results found no significant difference in adherence or other primary outcomes between participants receiving exercise instructions with the aid of an online tool compared with those receiving exercise instructions with paper-based methods. As the confidence intervals about the primary outcome of adherence were less than 34% (sample size estimation) the authors can therefore conclude that the use of the online prescription tool was not superior to traditional paper-based methods of prescribing exercises.

One concern is the number of apparent minor variations from the protocol in the outcome (e.g. proportion of days exercise) and in the analysis (e.g. planned incidence rate ratios for adherence) – see below for more detail. Although overall the trial has been completed according to the protocol it is important that any minor variations be declared and explained. Another concern was the fact that the analysis of the primary outcomes appears to have been completed as a per protocol analysis rather than as intention to treat as planned. For example, if there were log-book data of any participants withdrawing during the program their data should be included in the analysis.

General Comment Response: Thank you for the feedback. It was an oversight to not provide an explanation of the differences between the study procedure and the protocol, but we have rectified this now and included a new section in the Methods: "Deviations from protocol", whereby the explanations and justification for the changes have been outlined as such:

"In the course of implementing the study we made several decisions to deviate from the published protocol. Firstly, we decided to rely on the per protocol analysis because we did not have sufficient sample size to implement intention-to-treat methods such multiple imputations; furthermore, for 5 of the 8 participants who dropped out did we not have any adherence logbook data results to draw on. We reported an intention to treat analysis for the exploratory analysis of change in adherence across the 8-week intervention period, where we could use adherence logbook data from 3 participants with incomplete data sets. Secondly, we did not conduct the extra adherence statistical analysis, using Poisson regression, of the logbook adherence data that was proposed because the linear regression initiated as our first analysis accurately reflected the data.

Thirdly, our protocol specified that we would report on adherence using three methods: the number of exercise days completed, the number of exercises performed, and the number of repetitions

completed. It was decided to report on adherence using the latter two methods only, because reporting on the number of exercise days completed may be misleading for this study. In this study the home exercise programs were individualised in both the number of exercises performed per exercise day, and the number of exercise days per week. Furthermore, because we did not seek to take any steps to alter programs from normal clinical practice, for some participants there were variation in the number of exercises performed by one participant on a day to day basis; for example, some participants could have a stretching activity performed everyday as well as strengthening activities on 3 days per week. Given this broad variety in home program prescription, reporting on the number of exercises days was less specific than the other methods of reporting on adherence adopted, and potentially misleading, so it was not used.

Some specific comments
Comment 1. Abstract page 3, line 18: Stated that n= 54 were recruited, but numbers with CP and other disabilities only adds up to n=46.

Response 1: Thank you for bringing this to our attention. The breakdown by diagnosis in the abstract has been amended and now includes all 54 recruited participants (not just the 46 who completed the intervention).

“Results: Fifty-four participants with CP (n=37) or other neurodevelopmental disabilities (n=17) were recruited. Forty-six completed the 8-week program, with 24 the intervention group and 22 in the control.”

Comment 2. Abstract page 3, line 17: Please specify the primary and secondary outcomes separately.

Response 2: The outcomes have now been listed separately within the abstract:

“Primary outcome measures: Adherence to exercise programme, goal attainment, and exercise performance.
Secondary outcome measures: Enjoyment, confidence, and usability of Physitrack®.”

Comment 3. Abstract page 3, line 22: Clarify the sample size on which analysis was completed (n=46). It appears that a per protocol analysis has been reported, only including participants who completed the 8-week program.

Response 3: You are correct; a ‘per protocol’ analysis was used to generate the study outcomes. We did, however, complete an intention-to-treat analysis for the ‘change in adherence over time (week by week)’, because in this analysis we were able to access the available data provided from the participants who only had incomplete adherence data (n= 3), using mixed-models analysis. The specific number of participants incorporated in each analysis is stated in our CONSORT flow chart.

In the primary and secondary outcomes, however, the sample size was too small and there was insufficient data available to conduct an intention to treat analysis. Five of the 8 participants who dropped out did not have any logbook (or self-report) adherence data to draw on. We did not have a sample size large enough to identify certain participant characteristics that were associated with outcomes and use these in the multiple imputations to address missing data. Similarly, we did not think that the ‘last observation carried forward’ (LOCF) method would yield satisfactory results (e.g., LOCF assumption – see White et al., 2012).

White IR, Carpenter J, Horton NJ. Including all individuals is not enough: Lessons for intention-to-treat analysis. 9. 2012:396-407. doi:10.1177/1740774512450098

https://journals.sagepub.com/doi/10.1177/1740774512450098

In response to your queries here, we have now incorporated a new section in our Methods titled “Deviations from protocol” that includes this change from intention to treat analysis to a per protocol analysis. See “General Comment Response” above.

Comment 4. Strengths and limitations, page 4, line 15: Stated that recruitment target of n=58 was not attained (n=46). However, in the trial registry sample size estimation was n=66 including an allowance for attrition. Also, in the Abstract it stated that n=54 were recruited. It appears n=46 was the number completing the intervention.

Response 4: Thank you for highlighting this. The “Strengths and limitations” section has now been altered to clearly specify both the participant numbers recruited and the numbers completing the study, comparing for the power calculation numbers against the actual study numbers. This now provides greater clarity for the reader:

“Recruitment to the number specified in the power calculation (n=66) was not attained (n=54); similarly, the power calculation specified that 58 participants who completed the intervention (i.e. after dropouts) was required, yet only 46 participants completed the study.”

Comment 5. Data analysis page 9, line 24: According to protocol, adherence also measured as number of days exercised. This outcome has not been reported.

Response 5: You are correct, we had originally planned to report on adherence in three ways: exercise days, exercises performed, and number of repetitions completed. However, at the completion of the study we realised that reporting the adherence as the number of days complete would be misleading and inaccurate. We had individualised, rather than generic, programs; the number of exercise days/week were not standardised, so that each participant would differ in the number, sets, and repetitions of prescribed exercises. This approach to exercise prescription reflects common physiotherapy practice for children with disabilities; each child is provided a unique program to meet their individual goals and needs. In our study, the frequency of prescribed programs ranged from 2/week to 7/week. Furthermore, the number of exercises prescribed per day for some participants could be different across each week. For example, some participants would have a daily stretches, and also have strength exercises 3/week; hence, one participant may have 2 exercises on one day and 7 exercises on the next day.
Given such variability in the prescribed programs, we decided the "exercises performed" and "repetitions completed" would provide the most accurate analysis of adherence, and our focus in the adherence reporting. We have now added to our Methods at new section titled “Variations from Protocol” and have included this variation. See “General Comment Response” above.

Response 6. Data analysis page 9, about line 26: There appear to be some minor variations from the planned analysis of adherence. For example, incidence rate ratios were planned but not reported. Also, it was planned to adjust for number of follow up visits between the groups but this has not been reported. Also, it appears the analysis of weekly adherence, appropriately using a mixed effects model, was not planned - is that right?

Response 6:
Yes, as you have outlined, incident rate ratios were not reported as Poisson regression was not performed as originally planned as an additional analysis. Simple linear regression was used to analyse the logbook adherence data and proved suitable given the distribution of the data collected. This change has now been described in the “Variations from Protocol” section, as recorded in “General Comment Response” above.

We have now reported on the number of follow-up visits overall and in each group. This has been added to the results, in the section “Intervention characteristics”.

“We also considered the number of follow-up home visits or clinic visits during the 8-week intervention period from the 37 participants (of the 46 who completed the study) for which this data was available. Overall, the median number of follow-up appointments was 0 (IQR 0 to 2), with 53% having 0, 39% having 1-3, and 8% having 4-7. Comparing groups, the median number of follow-up appointments in the intervention group was 0 (IRQ 0 to 1) and the median in the control was also 0 (IQR 0 to 2).”

Yes, this is correct, the analysis of weekly adherence was not something we planned from the outset. Change in adherence over time, in each group, was a question of interest that emerged in the course of conducting the study. Hence, we described this as exploratory analysis on page 13, line 47.

Comment 7. Results page 10, about line 10: Please report the characteristics of participants who did not complete the intervention.

Response 7: These have now been added:

“The characteristics of the 8 participants who dropped out were heterogeneous in terms of age (mean 10.3, standard deviation (SD) 2.4), sex (female n=4, male n=4), and diagnosis (CP n=5, Autism Spectrum Disorder n=2, rare syndrome n=1).”

Comment 8. Results section: Results on primary outcomes were reported on 44, 46 and 40 participants, respectively. As presented, this presents as a per-protocol analysis rather than an intention to treat analysis, only including data for those who completed the intervention. If available, data from participants who dropped out during the program should be included.

Response 8: This issue has now been identified and addressed in the Methods section titled “Deviations from protocol”. See “Response 3” above.

Comment 9. Results: How many follow-up visits were there by the physiotherapists in each group – this does not appear to be reported.

Response 9: We have now reported on the number of follow-up visits overall and in each group. See “Response 6” above.

Comment 10. Results: What were the characteristics of typical exercise programs? E.g. How many exercises? Dose? Time to complete prescribed program. Were these factors similar between the groups?

Response 10: We have added an additional section to the results titled “Intervention characteristics” to provide this information:

“The home exercise programmes for both the intervention and control group were individualised by the treating physiotherapist to attain the goals identified by the family. Therapists could prescribe a number of exercises they deemed suitable. Across both groups the median number of exercises prescribed was 6 (Interquartile range (IQR) 5 to 8, minimum 2 and maximum 14). With regard to the number of exercises prescribed between groups, some variance was evident; the intervention group had a median of 6 (IRQ 6 to 7) and the control group had a median of 5 (IQR 4 to 8). Physiotherapists could also set the program frequency for the child; the median number of prescribed exercise days

per week was 3 (IRQ 3 to 5). Considering the groups separately, the intervention group (IQR 3 to 4) and the control group (IQR 3 to 5) were similar having a median of 3 exercise days per week."

Comment 11. Patient flow figure: Should include numbers analysed for exercise performance data

Response 11: The participant numbers for the quality of exercise performance analysis have now been added to Figure 1.

Comment 12. Discussion page 17: Can you propose why the intervention did not lead to hypothesised improvements in adherence and other primary outcomes? Could it be that the mode of exercise instruction is less important than the therapist who prescribes the exercises – as alluded to in the introduction. What are the clinical implications of the findings?

Response 12: Thank you for raising these questions; we think that addressing the proposed concepts has improved the depth of our Discussion. Included in the additions to the Discussion is the topic of the importance of the therapist and the interpersonal factors in supporting the family:

"…Despite efforts taken in this study to report on adherence as fully and accurately as possible, the primary method of measurement - exercise logbooks - may have contributed to the similar adherence results between groups particularly when previous work has indicated that exercise logbooks can themselves be used as a reinforcer of adherence. Similarly, the collection of exercise performance videos at 3 time-points (for COEP measure) may also have unintentionally reinforced adherence. The development of a valid measure of adherence to address the inconsistencies observed in the literature appears an important step to enable comparison of research findings for knowledge translation, as has also been recommended in other populations. A second potential reason for the similarities in adherence findings between groups is the intervention itself. Whilst Physitrack® provides features to support exercise programmes that are unavailable with conventional methods, it is not made specifically for children. A therapy prescription app designed for children with disabilities with features that will be engaging for them, such as a bright and playful colour scheme, games and rewards for completing exercises, may be more effective in improving program adherence in children than Physitrack®. Thirdly, it is important to consider that interpersonal factors including the therapist's ability to listen and partner with parents, and to provide ongoing follow-up support have been identified as key contributors to adherence and are beyond the delivery method investigated here. Accordingly, our finding in both groups of weekly reductions in adherence across the 8 weeks of the programme highlights the need in clinical practice to have regular follow-up to home exercise programmes and avoid a 'set and forget' approach."

Reviewer: 2
Reviewer Name: Robert Ware
Institution and Country: Griffith University, Australia
Please state any competing interests or state 'None declared': None declared

Thank you for the opportunity to review this interesting manuscript
I have some minor comments which I hope will improve readability.

Comment 1: Abstract/Results – between group difference for the primary outcome should be reported in abstract (in this case between-group mean difference (95%CI) for adherence)

Response 1: Thank you. This finding has been added to the Abstract/Results as recommended.

"There was no difference between the two groups for percentage of exercises completed (intervention (n=22): 62.8% (SD 27.5), control (n=22): 55.8% (SD 19.4), between group mean difference -7.0% (95% CI: -21.6 to 7.5, p=0.34))."

Comment 2: Abstract/Results – does "There was a downward trend in adherence in both groups of -2.3% (95% CI: -3.3 to -1.3, p<0.001) for each consecutive week" refer to both groups individually or both groups combined? – was there a between-group difference in rate of adherence over time? It should be noted in the abstract that this was an exploratory analysis (i.e. it was not pre-specified)

Response: 2 This sentence has now been removed from the abstract because it was an exploratory analysis, and also to remain within the 300-word limit whilst also implementing the other changes requested by editors and reviewers.

Comment 3: Abstract/Results (and throughout manuscript) - Usually p-values are reported to 2 decimal places to p=0.01, then 3 decimal places to P<0.001

Response 3: Thank you for this advice, the p value for adherence has been amended as you have recommended in the Abstract and through the manuscript.

Comment 4: Article summary – the article summary that follows the abstract would be more useful if it included the key results of the study

Response 4: We understand your recommendation and concur. However, we have not included this information because the BMJ Open Author guidelines for the Article Summary state: "An Article Summary, placed after the abstract, consisting of the heading 'Strengths and limitations of this study', and containing up to five short bullet points, no longer than one sentence each, that relate specifically to the methods. They should not include the results of the study." To our understanding, the publishers place the Strengths and Weaknesses adjacent to the Abstract, so the Abstract will provide some context for the reader around the study results when reviewing the strengths and weaknesses.

Comment 5: Methods – Methods section should be ordered according to the CONSORT guidelines – in particular the sample size section (item 7 in CONSORT 2010) should appear between Outcomes (item 6) and Randomisation (item 8)

Response 5: Thank you for identifying this point; the Methods section has been altered as you recommended. There is a new heading Sample Size which appears after Outcomes and before Randomisation; see page 8, line 56:

"Sample Size
Sample size was calculated using published data from a web-delivered intervention that measured adherence.30 To detect a difference in adherence of 85% in the intervention group and 51% in the control group, with at least 80% power and significance level of 0.05, we required 29 children per group (58 total), using a 2-tailed test. To allow for 15% attrition, we aimed to recruit 33 participants per group (66 in total)."

Comment 6: Methods – I was expecting to see an ethics statement regarding institutional approval towards the start of the Methods section.

Response 6: This has now been added:

“Ethical approval was obtained from Human Research Ethics Committee, Curtin University, Western Australia (ref number 10391). This study was prospectively registered with the Australian New Zealand Clinical Trials Registry (ACTRN12616000743460).”

Comment 7: Methods – ANZCTR trial registration number should be included in first paragraph of Methods

Response 7: This has now been added. See “Response 6” above.

Comment 8: Methods – how were children stratified according to the Functional Mobility Scale? What cut-off was used?

Response 8: This information has now been added to the Randomisation section:

“Participants were stratified by their level of functional mobility [as indicated by a rating of their mobility over a distance of 50m using the Functional Mobility Scale (FMS) (1= non-ambulant requiring wheelchair; 2-6= ambulant)] and by age (less than 12 years, or 12 years and older) for random allocation to one of two groups using a computerised random number generator.”

Comment 9: Methods/Data analysis – note that “Stata” is the name of the software rather than an acronym so should not be all upper-case

Response 9: Thank you for clarifying this. This change has now been made:

“All statistical analyses were performed using Stata version 15.1.”

Comment 10: Results/Table 1 – the Functional Mobility Scale and GMFCS (which is not defined) are listed in Table 1. Could these two scales be (briefly) defined in the Methods please? What do they measure and how are they ordered (e.g. does scoring a “1” mean the child has very good or very poor function)?

Response 10: Thank you for alerting us to this oversight. An explanation of the FMS and GMFCS has now been added to the footnote of Table 1 to provide the necessary context to this data:

“Functional Mobility Scale (FMS) is a tool for the classification of functional mobility in children, with a rating of 6 representing the most independently mobile, and a rating of 1 for children who are the least independently mobile and rely on wheeled mobility. FMS rates mobility at 3 distances: 5m, 50m and 500m; for the purposes of this study we chose to utilise the FMS to rate the participants’ mobility over a distance of 50m only. GMFCS: Gross Motor Functional Classification Scale-Expanded and Revised, is a functional mobility classification tool suitable for children with Cerebral Palsy. In the GMFCS “Level I” represents the most independently mobile through to “Level V” which represents the least mobile.”

Comment 11: Results/Adherence – rather than reporting the between group difference as “P>0.05” in the text, could the mean difference and 95%CI be reported instead please

Response 11: These details have now been added to the text:

“Adherence was approximately 60% in both groups. There was no statistically significant difference in the proportion of exercises performed [Between group mean difference -7.0 (95% Confidence Interval (CI) -21.6 to 7.5, p= 0.34)] or of exercise repetitions completed [Between group mean difference -8.6

(95% CI -23.8 to 6.5, p= 0.26)] between the intervention and control groups from our analysis of 44 participants whom provided complete adherence data: 22 in each group (see Table 2)."

Comment 12: Results/Table 2 – could mean difference (95%CI) be reported for the 'Adherence: Self-report" questions

Response 12: An extra column has been added to this part of Table 2 where the between group median difference results, along with their 95% CI's, have been added. See page 13, line 19-43.

Comment 13: Results/Adherence and goal attainment – could 95%Cis be reported for the estimates of the correlation coefficient please

Response 13: Thank you for identifying this. These 95% CI's have been added.

Exploratory analyses examined the degree to which goal attainment, as measured by the COPM, was related to adherence with the home exercise programme; the analysis revealed a small correlation between adherence and change in self-rated performance of goal activity (r=0.200, 95% CI: -0.103 to 0.469) and self-rated satisfaction in goal activity (r=0.050, 95% CI: -0.251 to 0.342).

Reviewer: 3
Reviewer Name: Georgina Clutterbuck
Institution and Country: Western Sydney University, Australia
Please state any competing interests or state 'None declared': None declared

Thank you for the opportunity to review your paper. The role of technology in exercise prescription and program adherence is an important area of research and your work has provided important groundwork for their application in the population of children with disabilities. Overall, I found the paper to be clear, concise and well designed, especially with the published protocol. I hope that the following feedback is helpful in strengthening the paper for publication.

Abstract: L13: Method- in the first sentence it reads like the community therapy was completed using Physitrack or paper-based methods, rather than the home exercise program.

Response 1: Thank you for clarifying this. The abstract Methods section has been amended to follow the BMJ Open structure, and the Abstract now includes the title "Intervention" clearly outlined as being the home program (rather than community therapy):

"Intervention: All participants completed an individualised home exercise program, which was delivered to the intervention group using Physitrack® and conventional paper-based methods for the control group."

Strengths and Limitations:
1. Concise and well communicated

Response 2: Thank you – we appreciate your positive feedback.

2. L17- It would be beneficial to name the tools here.

Response 3: These have now been added.

“Achievement of individualised goals was measured using the Canadian Occupational Performance Measure that has established reliability and validity evidence in past work, whereas the measure identified for evaluating the quality of exercise performance, Correctness of Exercise Performance scale, has not yet been formally evaluated.”

Introduction:
1. L15: It would be beneficial here to state that this definition is specifically for children with disabilities as it may differ from those devised for other population groups (i.e. parental participation would not be relevant for most adults).

Response 4: This has now been added into this sentence.

“Goal directed home programmes for children with CP encompass “therapeutic activities that the child performs with parental assistance in the home environment with the goal of achieving desired health outcomes”. “

2. L32: You have clearly identified how mHealth could be used to capitalise on the facilitators of adherence based on this research. It would also be beneficial to include some of the barriers to adherence that were identified in this research that mHealth could target, or which may be out of scope for Health.

Response 5: Thank you for this suggestion; it has led our team to go back through the qualitative studies on home programs with children with cerebral palsy. We have found that the studies really do emphasise the facilitators of effective home programs, based on the feedback of children and parents of children with disabilities. Whilst one of the papers described some negative experiences that parents experienced, these focussed on broader issues around the relationship with the physiotherapist, such as how therapists communicate with the parent during sessions and how they collaborate (or does not) with the parent in the planning of the program. These negative experiences speak to how a therapist should approach the development of home programs in a partnership, but not anything specific that could be dealt with in terms of technology. Given the strong focus on facilitators to effective home programs across this literature, we have removed the reference to barriers in this part of the Introduction:

“Qualitative research has investigated the facilitators of adherence to home programmes provided to children using conventional paper-based methods.”

Methods:
1. P8, L44: This sentence has some issues with plurality etc- a possible suggested edit: BoNTA injections were not an exclusion criterion because participation in an exercise program following BoNTA is recommended practice.

Response 6: Thank you for this suggestion. This sentence has been adjusted accordingly:

“Botulinum neurotoxin type A (BoNTA) injections were not an exclusion criterion because participation in an exercise programme following BoNTA is recommended practice.”

2. P9, L45: Did PTs use the library videos or customise their videos- do you have access to details about the proportion that used each type?

Response 7: As an observation whilst conducting the study, both custom and library videos were used; some therapists used only custom, some only library, and many a mixture of both. The custom video feature was used commonly and something that many therapists expressed enthusiasm about

when RJ conducted their training. However, in efforts to allay burden on our therapists we did not collect detailed data on this information.

3. Further details about the number of exercises and type of exercises (and if they were similar in both groups) would improve the report of the intervention.

Response 8: Details have been provided about the number of exercises prescribed and the program frequency, in a new section in the Results titled “Intervention characteristics”:

“The home exercise programmes for both the intervention and control group were individualised by the treating physiotherapist to attain the goals identified by the family. Therapists could prescribe a number of exercises they deemed suitable. Across both groups the median number of exercises prescribed was 6 (Interquartile range (IQR) 5 to 8, minimum 2 and maximum 14). With regard to the number of exercises prescribed between groups, some variance was evident; the intervention group had a median of 6 (IRQ 6 to 7) and the control group had a median of 5 (IQR 4 to 8). Physiotherapists could also set the program frequency for the child; the median number of prescribed exercise days per week was 3 (IRQ 3 to 5). Considering the groups separately, they were very similar with both the intervention group (IQR 3 to 4) and the control group (IQR 3 to 5) having a median of 3 exercise days per week.”

4. The protocol details that participants would receive weekly reminders, as well as continued intervention by the treating physiotherapist at the clinic or home. It would be beneficial to detail the number of home and centre visits received and if there was any difference between groups.

Response 9: Yes, the participants received weekly email or text reminders, primarily to ensure that we collected the weekly adherence log sheets. We have now reported on the number of follow-up visits overall and in each group:

“We also considered the number of follow-up home visits or clinic visits during the 8-week intervention period from the 37 participants (of the 46 who completed the study) for which this data was available. Overall, the median number of follow-up appointments was 0 (IQR 0 to 2), with 53% having 0, 39% having 1-3, and 8% having 4-7. Comparing groups, the median number of follow-up appointments in the intervention group was 0 (IRQ 0 to 1) and the median in the control was also 0 (IQR 0 to 2).”

5. P10, L14: When were the three time points of that the COEP recorded?

Response 10: This information has now been added:

“Primary outcomes included: adherence to the prescribed exercise programme (via weekly log book of exercise and repetition completion, and post-intervention participant responses to adherence questions using an 11-point Numeric Rating Scale (NRS)), achievement of individualised goals (rated using COPM before and after intervention), and performance of prescribed exercises (researcher assessed following viewing videos of participants performing prescribed exercises at three different time points: pre-intervention, mid-intervention, and immediately post-intervention, and scored using Correctness of Exercise Performance (COEP)).”

6. P10, L19: Extra comma after “we also measured”

Response 11: Thank you, this comma has been removed.

“Secondary outcomes included: enjoyment of exercise (using Physical Activity Enjoyment Scale (PACES) before and after the intervention), satisfaction with and confidence to complete programme

(using 11-point NRS26), and process measures (using 11-point NRS26). In the intervention group only, we also measured the usability of Physitrack® on a 5-point scale (using a modified System Usability Scale (SUS)).”

7. P11, L24: Adherence data were (not was)

Response 12: This has been altered as advised.

“Adherence data (variables: number of exercises, and number of repetitions completed) were calculated as a proportion of the prescribed program. We used mixed effects models to assess weekly adherence that accounted for correlations among repeated measures with time as a continuous covariate, and allowing for missing observations for the 3 participants who provided incomplete data (4, 5 and 6 weeks of log book adherence data, respectively).”

8. The journal has requested particular attention be paid to certain CONSORT items. One of these is allocation concealment- it is not clear if/how the number sequence was concealed after it was randomly generated so as to conceal the allocation of the next enrolled participant.

Response 13: In response to your query we have sought clarification on this form the CONSORT statement. This outlines that centralised or “third-party” assignment is a key feature of good concealment mechanisms. We have now specified this in our description of the randomisation process. Although the person conducting the concealed randomisation (Amanda Marie Blackmore) is a researcher known to the team, she was external to this study and not directly involved. She has expertise in health research processes and, hence, she was asked by the team to conduct the randomisation process following necessary methodology to maintain blinding, and also using block allocation methods whereby the allocation of the next enrolled participant was unknown to her until the process was conducted.

“Principal researchers were blinded to group allocation. Third-party assignment was used to conceal the participant allocation from the principal researchers. AMB implemented the randomisation process before contacting each treating physiotherapist directly by email to notify of group allocation, thereby maintaining blinding of researchers for the study duration. During data analysis, the principal researchers and biostatistician remained blinded to group allocation, with nominal group names assigned by AMB. It was impossible to blind the physiotherapists implementing the intervention, or the participants, to group allocation. However, physiotherapists and participants were blinded to the primary aims of the study until study completion.”

Results: Clearly reported.

Response 14: Thank you.

Discussion:
1. The paper would benefit from further discussion around the reasons that no difference between conventional and Physitrack adherence was found. In addition to the discussion around the responsiveness of the COEP, could the completion of the COEP (i.e. introducing filming of the home program to those who otherwise would only have a paper program) have influenced the feeling of being observed in the control group, thereby confounding the results? This may add to the comment regarding return of logbooks acting as an exercise reminder. There may have also been barriers to home exercise program adherence reported in the literature which continue with both methods.

Response 15: Thank you for this recommendation. We have made significant amendments to our Discussion based on this and similar comments from other reviewers. For the complete paragraph

see page 17, line 23 – page 18, line 7. Below is the discussion of potential reasons that no difference was found:

"…Despite efforts taken in this study to report on adherence as fully and accurately as possible, the primary method of measurement - exercise logbooks - may have contributed to the similar adherence results between groups particularly when a previous work has indicated that exercise logbooks can themselves be used as a reinforcer of adherence.4 Similarly, the collection of exercise performance videos at 3 time-points (for COEP measure) may also have unintentionally reinforced adherence. The development of a valid measure of adherence to address the inconsistencies observed in the literature appears an important step to enable comparison of research findings for knowledge translation, as has also been recommended in other populations.34 A second potential reason for the similarities in adherence findings between groups is the intervention itself. Whilst Physitrack® provides features to support exercise programmes that are unavailable with conventional methods, it is not made specifically for children. A therapy prescription app designed for children with disabilities with features that will be engaging for them, such as a bright colour scheme, games and rewards for completing exercises, may be more effective in improving program adherence in children than Physitrack®. Thirdly, it is important to consider that interpersonal factors including the therapist's ability to listen and partner with parents, and to provide ongoing follow-up support have been identified as key contributors to adherence and are beyond the delivery method investigated here. Accordingly, our finding in both groups of weekly reductions in adherence across the 8 weeks of the programme highlights the need in clinical practice to have regular follow-up to home exercise programmes and avoid a 'set and forget' approach."

2. Considering that 5% of participants were identified as potentially causing harm by doing their exercises, it may be appropriate to suggest that physiotherapists could use video to identify these children and therefore decrease harmful home exercise practice.

Response 16: This suggestion is excellent. This concept has now been outlined in the Discussion.

"…However, given the finding that 25% of the exercises reviewed were not performed correctly enough to achieve the exercise purpose and 5% of exercises were performed unsafely, we are able to recommend the use of video footage taken at home (e.g. by parents) to review how children perform prescribed exercises, in order to improve effectiveness and safety of home programmes in clinical physiotherapy practice."

3. While the results did not support the initial hypothesis that Physitrack would be superior, the outcome of being equally as effective as conventional paper-based programs indicates that therapists can confidently use these methods based on the individual preferences and circumstances of their patients.

Response 17: Thank you for identifying this. This has been added to our recommendations in the concluding paragraph of our Discussion:

"Conversely, neither did the measures yield worse results, or adverse findings, for participants using Physitrack®, hence the selection of programme delivery method (paper versus online) can remain at the discretion of the physiotherapist who is guided by the interests, needs and preferences of the child and parent with the disability."

Thank you again for the opportunity to review this paper. I look forward to seeing it published.

Reviewer: 4
Reviewer Name: Ferdinand von Walden
Institution and Country: Karolinska Institutet, Sweden
Please state any competing interests or state 'None declared': None declared

The authors have completed an ambitious project to investigate the adherence to and effectiveness of an 8-week home exercise program for children with disabilities delivered using Physitrack or paper-based methods. Although the study is technically correct and adheres to the CONSORT guidelines for presenting data and methods I have major concerns as the study is under-powered.

In the method section, the authors report the results of a power calculation that states the retirement of n=58 participants to be able to detect the expected difference between the two methods of program delivery. To allow for a few drop outs/non-finishers the aim was to recruit n=66. The actual number of participants that completed the trial was n=46. Given that no difference was detected between the programmes, the fact that the study is underpowered represents a problem in my view. Looking at the data, some data points suggest a better adherence in the intervention group (e.g. log book findings) that might have proved significant in a properly powered study. The fact that the study is underpowered to answer both primary and secondary outcomes makes me hesitant to recommend this manuscript for publication. The fact that the study is underpowered is not properly discussed in the discussion section.

Response: These issues are important, so we have sought to address them in our answers to your specific comments below. Namely, we have highlighted the extensive efforts that were taken to achieve the recruitment of n=46 that weren't previously recorded in the manuscript (see "Response 4" below). Also, we acknowledge that that study being underpowered was not sufficiently emphasised in the Discussion and this has now been corrected (see "Response 7" below). Collectively, therefore, we believe the results of this study are worth publishing because they provide qualitative (e.g., challenges with participant recruitment) and quantitative (e.g., data point for subsequent meta-analyses) contributions to the wider body of evidence in this area.

Minor comments:
Comment 1: I suggest that the authors should reconsider the title of the manuscript, as a large proportion of the participants were non-CP.

Response 1: Thank you. The target audience for this study was primarily children with cerebral palsy and we wanted this study to contribute to a continuum of research in the scientific literature on home programmes focussed on cerebral palsy. Conversely, we did not want to exclude children with other disabilities, and the nature of the research was around an intervention strategy (exercise programmes) that is frequently used by therapists and can be applied to a broader audience. This approach also harmonises with the clinical caseloads that we work with at Ability Centre, where Cerebral Palsy is the most common presentation, but we work with a range of other neurodevelopmental disabilities. We have taken on board your suggestion by broadening the title:

Can an online exercise prescription tool improve adherence to home exercise programmes in children with cerebral palsy and other neurodevelopmental disabilities?
A randomised controlled trial

Comment 2: Please make the number of study participants more clear in the abstract. It currently says fifty-four participants were recruited (n=32 CP and n=14 other). However the stated numbers (n=x) refers to the number of participants that completed the trial. The number of participants in the intervention group is not consistent (22, 24)

Response 2: We now clarified the participant numbers in the Results section of the Abstract, with greater specificity about the numbers that were recruited and the numbers that completed the study:

“Results: Fifty-four participants with CP (n=37) or other neurodevelopmental disabilities (n=17) were recruited. Forty-six completed the 8-week program, with 24 the intervention group and 22 in the control.”

And also, we have added detail to the Strengths and Weaknesses summary:

“Recruitment to the number specified in the power calculation (n=66) was not attained (n=54); similarly, the power calculation specified that 58 participants who completed the intervention (i.e. after dropouts) was required, yet only 46 participants completed the study.”

With regard to the sentence that records the intervention group having a sample of 22 participants, this number (and result) is in reference to the adherence findings, for which we had complete data sets for 22 of the 24 participants in the intervention group. This was further specified in the CONSORT flow chart (Figure 1, page 24): with participant numbers for each primary outcome analysis, also in the Results/Adherence section (page 12, line 35-36) and in Table 2 (page 13, line 9).

Participants:
Comment 4: What was the rational for mixing neurodevelopmental diagnosis with such variable motor component?

Response: 3 Thank you for your question. Part of the rationale for this decision is outlined in “Response 1” above. Home exercise programmes are a widely used intervention approach that is not limited to a particular diagnosis. The focus of this study was to investigate adherence within a paediatric cohort, as children are considered to have/be subjected to unique factors that affect adherence (such as the parent-child relationship) that are not common to adults. Similarly, having a neurodevelopmental disability introduces a particularly unique set of circumstances that will also impact on adherence, including children receiving life-long therapy, compared to typically developing children. However, we believe the factors influencing adherence to a home program are not unique to CP or any other particular neurodevelopmental diagnosis, and they can be included together in one study. In addition, the generalisability of results was important to the research team, as it aligns with the typically broad range of diagnoses that therapists work with in the disability sector in Australia.

Comment 4: It's stated in the methods section that "recruitment period was determined by funding timeline agreements". So the primary reason for stopping the study before it was sufficiently powered was financial?

Response 4: We apologise for this misdirection. Every effort was made to increase our recruitment for this study to achieve the participant numbers identified in the power calculation. Our initial 9-month recruitment and data collection period (recruitment occurred in an ongoing fashion, whilst data collection was occurring for earlier recruits) was extended to 21 months, for which we had to seek approval through the research funding body to extend the timelines of our grant and use the same grant funds over a longer period (with the primary researcher being paid at a lower rate to achieve this goal). In addition, we extended our recruitment beyond the primary location (Ability Centre) to other disability service providers, physiotherapy private practices, and government rural health centres within Western Australia. Through these strategies and many other ongoing efforts we were able to increase our recruitment to 80% of our power calculation. The only other option was to continue to extend the time frame of the study until more children within our recruitment catchment became old enough to meet the inclusion criteria; however, this was not a viable option within the (already extended) funding period:

"The recruitment period was determined by funding timeline agreements, and apparent exhausting of the recruitment catchment pool. Efforts to increase recruitment included extending the recruitment period from 9 months to 21 months and expanding our recruitment location from Ability Centre to other community physiotherapy providers in Western Australia."

Comment 5: Was prescribed oral medication documented?

Response 5: Prescribed oral medication was not recorded. We did record Botulinum Toxin Type-A injections, this being a common intervention in children with CP that requires therapy input, and we have added data about this intervention to Table 1. See page 11, line 49-52.

Comment 6: How many of the kids with CP received BTX during the intervention? Were these participants evenly distributed between the groups?

Response 6: This data has now been added to Table 1 - "Characteristics of the participants by group". See page 11, line 49-52.

Discussion:
Comment 7: The conclusion of the study is overly positive in my opinion. First, the study is underpowered to make a good interpretation of the data. Second, as is, the study indicates that the development of an app such as Physitrack seems unnecessary as the same outcomes can be obtained by paper-based documentation. Based on the available data, the discussion and conclusion is somewhat unbalanced.

Response 7: We have made amendments to the Discussion, and also the Abstract, to emphasise that the study was underpowered and hence that any conclusions here are preliminary. We have also recommended the need for a larger study.

The Abstract conclusion:

"Conclusion: Physitrack® provides a therapist with a new means of providing an exercise programme with online tools such as exercise videos, but our preliminary findings indicate that it may be no better than a traditional paper-based method for improving exercise adherence or the other outcomes measured. Exercise programmes remain an intervention supported by evidence, but a larger RCT is required to fully evaluate online delivery methods."

In the opening paragraph of the Discussion:

"The hypothesis that Physitrack®, which utilises features such as exercise videos, adherence tracking, and an app-based interface, would improve home programme adherence, quality of exercise performance, and goal attainment was not supported by the preliminary findings in this study. The method of delivery of the exercise programme employed here, in children with CP, did not appear to significantly affect adherence or the other outcomes. However, these findings need to be interpreted with caution given that we were unable to reach sufficient participation numbers in accordance with our power calculation."

And in the Discussion paragraph describing limitations of the study, this has been noted with greater clarity than before:

"Second, we did not achieve the recruitment of participants to the level specified in the power calculation, even with a 12-month extension and also expansion of our recruitment beyond Ability Centre to other therapy providers."

And in the concluding paragraph of the Discussion:

"However, given challenges with recruitment, these findings are preliminary in nature and a larger RCT is recommended to verify or refute these conclusions."

The app being proposed is very different to Physitrack, in particular, the way the app functions to engage with children as primary users. Following the conclusion of the study, some informal feedback from families and physiotherapists involved with the study indicated that Physitrack® did not address the needs and interests of children, and this may be Physitrack's primary downfall for the children with CP and other disabilities. They indicated they were looking for fun and interactive features, and in-app rewards for completing exercises. Literature in mobile health applications indicates that the execution of the app function and interface is key to its success in achieving health outcomes. Whist Physitrack has a smooth interface for general users, and a broad range of tools, it does not execute the functionality in a way that will optimise the interest and engagement of children. As noted, it does not make use of recognised behaviour change theory and strategies for altering health-related behaviour, nor does it incorporate gamification elements. Employing these features, recognised in the literature, would make for a very different app with potentially very different outcomes, and hence this is a recommendation from our experience of implementing this study.

VERSION 2 – REVIEW

REVIEWER	Nicholas Taylor La Trobe University, Australia I am co-author with Natasha Bear on another manuscript submitted for consideration for publication.
REVIEW RETURNED	12-Oct-2020

GENERAL COMMENTS	The authors have completed a thorough revision resulting in an improved report of their trial. I have just a couple of minor comments. 1. The flow chart says n=53 were randomized, Elsewhere, e.g. Abstract it is is n=54. Please correct. 2. Minor but on line 18, should goal attainment be changed to goal achievement to be consistent with the rest of the report. 3. p47, line 53: replace 'made' with 'designed'.

REVIEWER	Robert S Ware Griffith University, Australia
REVIEW RETURNED	02-Oct-2020

GENERAL COMMENTS	Thank you to the Authors for the thorough revision of their manuscript. My only minor comment is that the P-values in Tables 2 and 3 should be reported to 2 decimal places.

REVIEWER	Georgina Clutterbuck Western Sydney University
REVIEW RETURNED	08-Oct-2020

GENERAL COMMENTS	Thank you very much for the revision of this manuscript. The additional information you have added has strengthened it considerably and I hope to see it published shortly.

REVIEWER	Ferdinand von Walden Karolinska Institutet, Stockholm, Sweden
REVIEW RETURNED	01-Oct-2020

GENERAL COMMENTS	The authors have made considerable changes to the manuscript in accordance with the recommendations from reviewers and the manuscript is in my opinion much improved and more balanced. I don't have any further suggestions and think that the manuscript is fit for publication.

VERSION 2 – AUTHOR RESPONSE

Response to Reviewers

Reviewer: 1

Reviewer Name: Nicholas Taylor

Institution and Country: La Trobe University, Australia

Please state any competing interests or state 'None declared': I have recently been a co-author on another manuscript submitted for consideration for publication with author Natasha Bear.

The authors have completed a thorough revision resulting in an improved report of their trial.

I have just a couple of minor comments.

Response: Thank you for the feedback that facilitated an improved manuscript.

Comment 1. The flow chart says n=53 were randomized, Elsewhere, e.g. Abstract it is is n=54. Please correct.

Response 1: The CONSORT chart is correct that 53 children were randomised: this is after 54 participants were initially recruited, but then 1 early withdrawal occurred (as noted on the upper right-hand side of the chart). We do agree, however, that this creates some lack of clarity with the manuscript referring the 54 people recruited. To address this, we have added a sentence to the abstract to be more specific with the numbers randomised.

Results: Fifty-four participants with CP (n=37) or other neurodevelopmental disabilities (n=17) were recruited. Fifty-three were randomised after 1 early withdrawal. Forty-six completed the 8-week program, with 24 in the intervention group and 22 in the control.

Comment 2. Minor but on line 18, should goal attainment be changed to goal achievement to be consistent with the rest of the report.

Response 2: Thank you for highlighting this in the abstract. This change has been made.

Primary outcome measures: Adherence to exercise programme, goal achievement, and exercise performance.

We also noted the use of “goal attainment” in a number of points in the body of the manuscript. These have also been changed to “goal achievement” for consistency.

Comment 3. p47, line 53: replace 'made' with 'designed'.

Response 3: Thank you for noting this issue with phrasing. This has been corrected.

Whilst Physitrack® provides features to support exercise programmes that are unavailable with conventional methods, it is not designed specifically for children.

Reviewer: 2

Reviewer Name: Robert Ware

Institution and Country: Griffith University, Australia

Please state any competing interests or state ‘None declared’: None declared

Thank you to the Authors for the thorough revision of their manuscript.

Response: Thank you for your guidance with the revision of this paper, particularly with the statistical analysis and reporting.

Comment 1. My only minor comment is that the P-values in Tables 2 and 3 should be reported to 2 decimal places.

Response 1: Tables 2 and 3 have now been altered so that the P-values are reported to 2 decimal places instead of 3.

Reviewer: 3

Reviewer Name: Georgina Clutterbuck

Institution and Country: Western Sydney University, Australia

Please state any competing interests or state ‘None declared’: None declared

Thank you very much for the revision of this manuscript. The additional information you have added has strengthened it considerably and I hope to see it published shortly.

Response: Thank you for your feedback and your input in the development of this manuscript.

Reviewer: 4

Reviewer Name: Ferdinand von Walden

Institution and Country: Karolinska Institutet, Sweden

Please state any competing interests or state ‘None declared’: None declared

The authors have made considerable changes to the manuscript in accordance with the recommendations from reviewers and the manuscript is in my opinion much improved and more balanced. I don't have any further suggestions and think that the manuscript is fit for publication.

Response: Thank you for your recommendations and the opportunity to improve this paper.