E-HEALTH "STANDINGTALL" BALANCE EXERCISE FOR FALL PREVENTION IN OLDER PEOPLE: RESULTS OF A TWO-YEAR RANDOMISED CONTROLLED TRIAL

Journal:	*BMJ*
Manuscript ID	BMJ-2020-063155
Article Type:	Research
BMJ Journal:	BMJ
Date Submitted by the Author:	18-Nov-2020
Complete List of Authors:	Delbaere, Kim; University of New South Wales, Neuroscience Research Australia; University of New South Wales, School of Population Health Valenzuela, Trinidad; The University of Sydney, Physical Activity, Lifestyle, Ageing and Wellbeing Faculty Research Group, Faculty of Medicine and Health; Universidad Finis Terrae, Exercise Science Laboratory, School of Kinesiology Lord, Stephen; Neuroscience Research Australia, Falls and Balance Research Group; University of New South Wales, Neuroscience Research Australia Clemson, Lindy; University of Sydney, Faculty of Health Sciences Zijlstra, Gertrud; Maastricht University/CAPHRI, Health Services Research Close, Jacqueline; University of New South Wales, Neuroscience Research Australia; University of New South Wales, Prince of Wales Hospital Clinical School Wai-Chun Lung, Thomas ; The George Institute for Global Health; The University of Sydney Faculty of Medicine and Health, School of Public Health Woodbury, Ashley; University of New South Wales, Neuroscience Research Australia Chow, Jessica; University of New South Wales, Neuroscience Research Australia Miles, Lillian; University of New South Wales, Neuroscience Research Australia McInerney, Garth; University of New South Wales, Neuroscience Research Australia Toson, Barbara; University of New South Wales, Neuroscience Research Australia; Flinders University College of Medicine and Public Health Briggs, Nancy; University of New South Wales, Stats Central, Mark Wainwright Analytical Centre van Schooten, Kimberley; University of New South Wales, Neuroscience Research Australia; University of New South Wales, School of Population Health
Keywords:	accidental falls, balance control, ageing, exercise, ehealth, aged, 65 and over

SCHOLARONE™
Manuscripts

E-HEALTH "STANDINGTALL" BALANCE EXERCISE FOR FALL PREVENTION IN OLDER PEOPLE: RESULTS OF A TWO-YEAR RANDOMISED CONTROLLED TRIAL

Kim Delbaere (0000-0002-5655-0234), Trinidad Valenzuela, Stephen R. Lord (0000-0002-7111-8802), Lindy Clemson (0000-0003-2687-1114), G.A. Rixt Zijlstra, Jacqueline C.T. Close, Thomas Lung (0000-0001-9978-6311), Ashley Woodbury, Jessica Chow, Garth McInerney, Lillian Miles, Barbara Toson (0000-0001-6661-9971), Nancy Briggs (0000-0002-1134-0807), Kimberley S. van Schooten (0000-0003-0902-8440)

Kim Delbaere
Associate Professor
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia. School of Population Health, University of New South Wales, Kensington, 2052, Australia

Trinidad Valenzuela
Postdoctoral Research Associate
Physical Activity, Lifestyle, Ageing and Wellbeing Faculty Research Group, Faculty of Medicine and Health, The University of Sydney, Sydney, 2006, Australia. Exercise Science Laboratory, School of Kinesiology, Faculty of Medicine, Universidad Finis Terrae, Santiago, 7501015, Chile

Stephen R. Lord
Professor
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia. School of Population Health, University of New South Wales, Kensington, 2052, Australia

Lindy Clemson
Professor
Faculty of Medicine & Health, The University of Sydney, Sydney, 2006, Australia

G.A. Rixt Zijlstra
Professor
Care and Public Health Research Institute, Department of Health Services Research, Maastricht University, Maastricht, 6229GT, The Netherlands

Jacqueline C.T. Close
Professor
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia. Prince of Wales Hospital Clinical School, University of New South Wales, Kensington, 2052, Australia

Thomas Lung
Senior Research Fellow
The George Institute for Global Health, University of New South Wales, Newtown, 2042, Australia. Faculty of Medicine and Health, School of Public Health, The University of Sydney, Sydney, 2006, Australia

Ashley Woodbury
Research Assistant
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia.

Jessica Chow
Research Assistant
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia.

Garth McInerney
Research Assistant
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia.

Lillian Miles
Research Assistant
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia.

Barbara Toson
Senior Research Fellow
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia. College of Medicine and Public Health, Flinders University, Adelaide, 5042, Australia

Nancy Briggs
Senior Statistical Consultant
Stats Central, Mark Wainwright Analytical Centre, University of New South Wales, Kensington, 2052, Australia

Kimberley S. van Schooten
Postdoctoral Fellow
Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, 2031, Australia. School of Population Health, University of New South Wales, Kensington, 2052, Australia

Correspondence to: Kim Delbaere k.delbaere@neura.edu.au

SUMMARY

Objectives: The frequency and costs of falls result in an increasing need for effective, self-managed fall prevention programmes for older people. The objective of this study was to test whether *StandingTall,* an unsupervised, home-based, e-Health balance exercise programme delivered via an App, could fill this void and prevent falls in community-dwelling older people.

Design: Assessor-blinded randomised controlled trial

Setting: Older people living independently in the community

Participants: 503 individuals aged 70+ who were independent in activities of daily living, without cognitive impairment, progressive neurological disease or any other unstable or acute medical condition precluding exercise.

Interventions: Participants were block randomised to an intervention group (IG: 2 hours of *StandingTall* per week + health education; N=254) or a control group (CG: health education; N=249) for 2-years.

Main outcome measures: The primary outcomes were rate of falls and number of fallers over 12-months. Secondary outcomes included falls, injurious falls, adherence, mood, health-related quality of life and activity levels over 2-years, and balance and mobility outcomes over 12-months.

Results: Both groups had a similar rate of falls and proportion of fallers at 12-months (p=0.071 and p=0.461 respectively). The IG had a 16% lower rate of falls over 2-years compared to CG (incidence rate ratio:0.84, 95% confidence interval, 95%CI:0.72-0.98, p=0.027). Both groups had a similar proportion of fallers at 2-years (p=0.239), but the proportion of injurious fallers at 2-years was 20% lower in IG compared to CG (relative risk:0.80, 95%CI:0.66-0.98, p=0.031). In the IG, 68.1% and 52.0% of participants exercised up to 12-months and 2-years for a median of 114.0 (interquartile range, IQR:53.5) and 120.4 (IQR:38.6) weekly minutes respectively. Groups remained similar in mood and activity levels. The IG had a 0.03

(95%CI:0.01-0.06) improvement on the EQ-5D-5L utility score at 6-months, and an improvement in standing balance of 11-seconds (95%CI:2-19) at 6-months and 10-seconds (95%CI:1-19) at 12-months. No serious training-related adverse events occurred.

Conclusions: *StandingTall* balance exercise did not significantly affect our primary outcomes. It did significantly reduce the rate of falls and number of injurious fallers over 2-years. e-Health exercise programmes may be a promising scalable fall prevention strategy.

Trial registration: ACTRN12615000138583

Funding: Australian National Health and Medical Research Council, Philanthropy.

WHAT IS KNOWN (2-3 sentences)

- Balance exercise programmes are amongst the most effective fall prevention strategies, with fall reduction rates of 23% in older community-dwelling people.
- Despite strong evidence that falls can be prevented, sustained full adherence in effective trials is poor with pooled adherence rates of 21% (range 0-68%) at 12-months.
- Previous studies have shown that e-Health technology can offer a well-accepted method for delivering unsupervised balance exercise to older people, with good adherence rates.

WHAT THIS STUDY ADDS (1-2 sentences)

- This study is the first large, long-term, and methodologically robust trial to examine unsupervised technology-driven exercise as a strategy to prevent falls in older people.
- The *StandingTall* programme did not significantly affect rate of falls and proportion of fallers at 1-year; however, *StandingTall* did significantly reduce the rate of falls and number of injurious fallers over 2-years with a dose adherence of 30 to 40%.

INTRODUCTION

Falls and fall-related injuries have persisted over the past three decades as a leading cause of morbidity and mortality in older people(1). With a rapidly ageing population globally, sustainable access to evidence-based cost-effective fall prevention programmes is a priority. High-quality systematic review and meta-regression evidence has confirmed that well-designed exercise programmes are amongst the most effective fall prevention strategies for community-dwelling older people, with fall reduction rates averaging 23%(2). However, to achieve similar effectiveness at a population level, we need a programme that people can access easily and adhere to long-term. Previous studies have found that older people prefer home-based exercises and that the inclusion of balance exercises is associated with higher adherence(3). Nevertheless, sustained adherence to prescribed home exercise programmes is low with pooled estimates of 21% (range 0-68%)(4). Studies providing a physiotherapist-led programme and/or a moderate level of home visits (i.e. <1 home visit per month and >2 home visits in total) achieve higher levels of adherence(4), however, this substantially increases the cost and reduces the feasibility as a population approach.

Digital technology can provide engaging and widely accessible methods for delivery of exercise programmes to enhance long-term motivation and adherence at relatively low cost(5). However, the provision of a well-designed unsupervised exercise programme, that is tailored and progressive in nature, yet safe, could be a challenge. *StandingTall* is a unsupervised, home-based e-Health balance exercise programme provided via an App, developed using principles of consumer design to ensure an appropriate and user-friendly interface for older people. Behavioural change strategies are incorporated to enhance exercise uptake and long-term adherence(6).

This randomised controlled trial aimed to determine the effect of *StandingTall* on the recommended set of core outcomes for fall prevention trials in older people (i.e. fall rate, number of fallers and injurious fallers, and known fall risk factors including balance, gait, concern about falling, health-related quality of life and physical activity levels)(7) over a 24-month follow-up period, when compared with a health promotion education 'control' programme.

METHODS

Study design

We conducted a prospective, assessor-blinded, two-arm parallel randomised controlled trial with 2-year follow-up in Sydney, Australia. The trial was approved by the UNSW Ethics Committee in December 2014 (HC#14/266). It was registered prospectively in the Australian and New Zealand Clinical Trials Registry (ACTRN12615000138583) on 13 February 2015 and the study protocol was published in 2015(6). The statistical analysis plan was preregistered in October 2018 via the OpenScience framework (https://osf.io/42gje/) prior to completion of data collection in November 2019. We used the CONSORT statement, ICMJE recommendations and TiDieR checklist for preparation of this manuscript.

Patient and Public Involvement

StandingTall was developed using consumer design principles. A group of older people were involved during the development of the *StandingTall* Application. They were asked to evaluate an early version on its usability and age appropriateness as a means to engage in fall prevention exercises using tablet-based technology. There was no other formal patient and public involvement in this study.

Participants

We recruited community-living older people in the Sydney metropolitan area via flyers, advertisements in local newspapers, presentations at residential and community senior centres, and word of mouth. After initial screening by telephone, eligible individuals were invited to participate if they were aged 70+ years, living in the community, independent in activities of daily living, able to walk household distances without the use of a walking aid, and willing and able to give informed consent and comply with the study protocol. Individuals were excluded if they had an unstable or acute medical condition that precluded exercise participation, suffered from a progressive neurological condition (such as Parkinson's disease, Multiple Sclerosis), were cognitively impaired as defined by a Pfeiffer Short Portable Mental Status Questionnaire (SPMSQ) score <8(9), or were currently participating in a fall prevention programme. Eligibility was determined after informed verbal consent. Eligible individuals who agreed to participate in the study were asked to provide informed written consent.

Randomisation and masking

Participants were randomised after completion of the baseline assessment. Permuted block randomisation with mixed block lengths of four and six was applied to form two groups of similar size (allocation ratio 1:1). People living in the same household were treated as one unit to avoid contamination. Allocation was performed centrally using a web-based randomisation programme by an investigator not involved in participant assessments or provision of the intervention. Allocation concealment was ensured as the randomisation code was only released after the baseline assessment was completed. Outcome assessors were blinded to study group assignment throughout the trial. Statistical analyses were performed blinded for intervention or control group allocation.

Procedures

All participants received a tablet computer with a health promotion education programme which focused on health-related information relevant to older people, in addition to usual care, for two years. This health promotion education programme comprised weekly fact sheets (104 in total) with information on healthy diet, medications, fall risk factors and exercise. The tablet-based health education alone was chosen as the active control intervention to control for the use of technology and allow data collection (i.e. falls during the trial period) through a tablet computer for both groups.

The intervention group received the *StandingTall* programme, with exercise equipment (foam cushion, stepping box, exercise mat), in addition to the health promotion education programme and usual care. This intervention consisted of balance exercises delivered through a tablet computer in the participants' homes with embedded behavioural change techniques, including a weekly calendar for scheduling exercises, goal setting and educational fact sheets. Participants were asked to exercise for at least two hours per week for the duration of the trial. The intervention was introduced gradually; participants commenced with 40-minutes of exercise per week, which was increased by 20-minutes fortnightly until participants reached the required dose of 120-minutes per week in week 9. *StandingTall* delivers individually-tailored balance exercises that increase in difficulty over time. It also allows people to choose the time and duration of their exercise sessions. The intensity of the balance exercises is monitored using a self-report modified rate of perceived exertion scale and is adjusted as performance changes throughout the trial without the need of supervision. Exercise adherence (volume, frequency) was monitored for 2-years following automatic data transfer to a server. During the first 6-months, participants were encouraged to inform the research team when they were going away or would not be able to exercise for a few weeks. Participants who did not inform the team and did not reach 100-minutes for two consecutive weeks were contacted by telephone to record the reason for non-adherence, discuss any issues related to the programme and encourage adherence. These calls stopped after 6-months to gain a better understanding of behavioural change and long-term exercise adherence.

Intervention group participants (IG) received two home visits. At the first visit, a qualified exercise physiologist instructed the participant on how to use the *StandingTall* programme. This occurred between 1 and 3-weeks after the baseline assessment and lasted approximately 1-hour. The second home visit of approximately 30-minutes at 1-month ensured safe use and progression of training. Control group participants (CG) received two phone calls by qualified exercise physiologists at the same time points, to discuss any issues with accessing the health education programme and using additional features of the tablet computer.

Outcomes

The *primary outcome measures* were the rate of falls and proportion of fallers over the first 12-months of the trial. A fall was defined as "an unexpected event in which the participant comes to rest on the ground, floor or lower level"(7). Falls were monitored using prospective weekly fall diaries through the tablet computer (over 24-months after baseline assessment). Fall information was automatically uploaded to a database. Research staff contacted participants with missing falls diaries by telephone at the end of each month to complete missing data. The falls database was checked, reviewed and locked before group allocation was unmasked. Falls that occurred up to 365-days (1-year) after randomisation were included in the primary analysis. Falls that occurred up to 730-days (2-years) after randomisation were included as secondary falls outcomes. Injurious falls were defined as falls resulting in any injury (e.g.

bruises, cuts/grazes, joint dislocations, sprains/strains, fractures, pain), or falls that required medical care (e.g. visit to physician, emergency department).

Secondary outcome measures were assessed at baseline, at 6-months to examine acute effects, and at 12, 18, and 24-months to examine retention effects. These measures included lab-based balance and neuropsychological assessments (at baseline and at 6 and 12-months after baseline assessment in the first 226 participants), and remote at-home measures of wellbeing, quality of life and activity levels (at baseline and at 6, 12, 18, and 24-months after baseline assessment in all participants). Physiological fall risk was assessed using the Physiological Profile Assessment (PPA)(10). Balance, functional mobility and gait were assessed using tests of standing balance (standing with feet in different positions for a maximum of 30-seconds per condition: feet together, near-tandem, and tandem on floor and foam cushion and on left and right foot on floor; sum of durations for all eight conditions), maximum forward-backward and controlled leaning balance,(10) timed sit-to-stand(11), and up-and-go tests(12), short physical performance battery(13), and self-selected walking speed over 10-meters(14). Stepping performance was assessed with Choice, Stroop and Inhibitory Stepping Reaction Time tests (15, 16). Cognitive function was assessed with the Montreal Cognitive Assessment(17), Trail-Making Tests (TMT)(18), and the Victoria Stroop task(19). Psychological outcome measures were assessed using by the iconographical Falls Efficacy Scale (concern about falling)(20), nine-item Patient Health Questionnaire (mood)(21) and the COMPAS-W scale (wellbeing)(22). Health-related quality of life was assessed using the 12-item WHO Disability Assessment Schedule(23), 5-level EuroQol- 5 Dimension (EQ-5D-5L)(24), and 20-item Assessment of Quality of Life 6-Dimensions (AQoL-6D) questionnaires(25). Detailed self-report information on frequency and duration of physical activity was assessed using the Incidental and Planned Exercise Questionnaire (IPEQ)(26). Daily-life activity was assessed with the McRoberts MoveMonitor (McRoberts, the Netherlands) as the average duration of daily walking and standing, and number of walking and standing bouts per day(27). A bout was defined as a period of consecutive activity. Because participants were instructed to remove the device before going to bed, we required a minimum wear duration of 12-hours per day on one or more days for daily-life activity data to be included in the analysis. Daily-life activity data were collected over a median of 6 (IQR 1) days for both groups.

Process outcome measures were captured as exercise duration via the tablet computer. Since participants were allowed exercise breaks when they were sick or went on holiday, we averaged weekly exercise duration as median values within individuals as a robust measure of central tendency. We obtained subjective user experience by assessing usability, enjoyment and exercise self-efficacy with the System Usability Scale (SUS)(28), Physical Activity Enjoyment Scale (PACES)(29), Exercise Self-Efficacy Scale (ESES)(30), and Attitudes to Falls-Related Interventions Scale (AFRIS)(31).

All outcome measures were assessed by trained exercise physiologists or physiotherapists blinded to group allocation. Safety was assessed in terms of adverse events, which were defined as any fall related to the prescribed exercise programme or involving the intervention equipment.

Statistical analysis

Sample size calculation. Based on previous evidence, we carried out an a-priori sample size calculation (5% significance level, 80% power, 33% effect, 20% dropout rate). This indicated that a sample of 500 would be sufficient to evaluate the efficacy of the intervention on the rate of falls and that a subsample of 200 would be sufficient to detect between-group differences in physical outcome measure changes(6).

Analysis plan. Analyses were conducted according to the pre-defined statistical analysis plan, as registered on the Open Science Framework (osf.io/42gje/). Data were coded to maintain group allocation blinding during analysis. Effectiveness analyses of the primary outcome were conducted on an intention-to-treat basis by a statistician (BT/NB) and independently replicated by one of the investigators (KSvS). The alpha level was set to 5%. Analyses were performed with Stata (version 16, Stata Corp.) and SPSS (version 25, IBM Corp.).

Missing data. In line with intention-to-treat principles, participants who were randomly assigned to a group were included in the analysis irrespective to their level of compliance with their group assignment. The primary outcome measures (i.e. number of falls per person-year and proportion of fallers over 12-months) were analysed without imputation or adjustment for descriptive characteristics, and with correction for follow-up duration when appropriate. The faller status of people with incomplete follow up (n=66 at 12-months and n=188 at 24-months) was assumed to be maintained during censoring. We used Little's MCAR test to determine the missing data patterns of secondary outcome measures. The secondary outcome measures were imputed using estimated means single imputation if they were Missing Completely At Random (MCAR), or under the assumption of Missing At Random (MAR) using multiple imputation to create 20 imputation datasets under joint multivariate normal imputation if they were not MCAR. Psychological wellbeing, health-related quality of life and physical activity questionnaire data were missing for 58 out of a total of 503 people at 6-months, for 82 people at 12-months, for 98 at 18-months and for 99 people at 24-months. Daily-life activity monitoring data was unavailable for 21 at baseline, 101 people at 6-months, 138 people at 12-

months, 148 people at 18-months and 156 people at 24-months. Clinic-based balance and neuropsychological assessment data were missing for 42 people at 6-months, and for 47 people at 12-months. These data were missing because of dropout, scheduling issues, non-adherence or technical problems. Little's MCAR test indicated that all data were missing at random with respect to participant baseline characteristics.

Primary outcomes. Primary outcomes were: *(i)* the number of falls per person-year, and *(ii)* the proportion of fallers over 12-months. The number of falls per person-year was analysed using Poisson regression to estimate the difference in fall rates between the two groups. The incidence rate ratio and its 95% confidence interval (95%CI) are reported. Poisson regression was selected over negative binomial regression (as a-priori registered in our statistical analysis plan) to allow for a direct comparison to our planned complier average causal effects analysis since the latter was based on a Poisson model. The results for the negative binomial regression can be found in *Appendix 1*. Days of follow-up was included as an exposure term in these models, i.e. the natural logarithm of the days of follow-up was added as an offset. The proportion of fallers in the two groups was examined using modified Poisson regression models for binary outcomes. Faller status was compared using: 0 falls versus 1+ falls; and relative risks and their 95% CIs are reported.

Secondary outcomes. Secondary fall outcomes were the number of falls, the complier averaged causal effect, proportion of fallers, and proportion of injurious fallers at 2-years. We employed instrumental variable regression to correct for imperfect participant adherence and gain insight into efficacy by estimating the complier averaged causal effect (CACE). We used a 2000-times bootstrapped 2-stage CACE estimator comprised of a linear regression with adherence as the dependent variable and group as the independent variable to obtain an estimate for adherence, followed by a robust Poisson regression with falls as the dependent variable, the natural logarithm of follow up in days as exposure, to estimate the effect of the intervention among people with perfect adherence. The number of injurious falls per person-year was analysed using Poisson regression to estimate the difference in injurious fall rates between the two groups. We analysed secondary non-fall outcome measures with robust generalised linear models using an exchangeable working correlation matrix and compared the change in scores over time at 6, 12, 18 and 24-months between IG and CG. When the residuals of the generalised linear models deviated from normality, we used a 1000-times bootstrap for each imputation dataset to obtain confidence intervals.

RESULTS

Between February 2015 and October 2017, 823 individuals were screened (*Figure 1*). Five hundred three people were included in the study and randomly assigned to IG (n=254) or CG (n=249). We lost 90 participants during the 2-year follow-up (n=53 in IG and n=37 in CG) and 46 IG participants discontinued the intervention but kept contributing data. Baseline characteristics of all participants are provided in *Table 1*.

Effect on primary fall outcomes

Rate of falls at 12-months. The average rate of falls over the first 12-months was 0.60 (SD 1.05) falls per year in IG and 0.76 (SD 1.25) in CG. The difference in fall rate was not statistically different, with an incidence rate ratio (IRR) of 0.82 (95% confidence interval; 95% CI 0.66-1.02), p=0.071, in IG compared to CG (see *Figure 2*).

Proportion of fallers at 12-months. Overall, 188 participants (37.4%) fell at least once in the 12-month follow-up. IG and CG participants were equally likely to fall at least once, with a relative risk of 0.90 (95% CI 0.67-1.20), *p*=0.461, in the IG compared to the CG.

Effect on secondary fall outcomes

Rate of falls at 24-months. The average rate of falls over the two-year follow-up was 0.57 (SD 0.95) falls per year in IG and 0.72 (SD 1.17) in CG. The difference in fall rate was statistically different, with an IRR of 0.84 (95% CI 0.72-0.98), *p*=0.027, in IG compared to CG.

Rate of falls via complier averaged causal effects at 24-months. CACE analysis revealed an IRR of 0.72 (95% CI 0.21-1.13), *p*=0.324, in IG compared to CG. This IRR was similar to that of the intention-to-treat analysis.

Proportion of fallers at 24-months. 270 participants (53.7%) fell at least once in the 24-month follow-up. IG and CG participants were equally likely to fall at least once, with a relative risk of 0.87 (95% CI 0.68-1.10), *p*=0.239), in the IG compared to the CG.

Proportion of injurious fallers at 24-months. 210 participants (41.7%) experienced an injurious fall during the 24-month follow-up. IG participants were less likely to be injurious fallers than CG participants, with a relative risk of 0.80 (95% CI 0.66-0.98, *p*=0.031).

Effect on secondary outcomes on wellbeing, quality of life and activity levels

We found no significant difference in psychological wellbeing or physical activity levels at 6, 12, 18 and 24-months in IG compared to CG (*Table 2*). We did find a small improvement of 0.03 (95% CI 0.01-0.06) on the EQ-5D-5L utility score at 6-months in IG compared to CG. All other health-related quality of life measures showed no difference between IG and CG groups at all time points.

Effect on secondary outcomes of balance and neuropsychological assessments

The 226 participants (45% of the total sample; 114 IG and 112 CG) who were invited for laboratory re-assessments were on average 1.1 years older (78.0 years (SD 5.4) vs 76.9 years (5.5) in those not invited; t(501)=-2.294, p=0.022) and scored 0.40 points higher on physiological fall risk (measured with PPA (1.10 (SD 0.82) vs 0.70 (0.90) in those not invited; t(501)=-5.063, p<0.001). There were no other significant differences in baseline characteristics between these groups. We observed a significant improvement in standing balance at 6 and 12-months (11-seconds, 95%CI 3-19 and 10-seconds, 95%CI 1-19, respectively) in IG compared to CG (*Table 3*). We found no significant difference in physiological fall risk, maximum forward-backward and controlled leaning balance, functional mobility and gait tests, stepping performance, or cognitive and executive functions at 6 or 12-months in IG compared to CG.

Subgroup analyses

Planned subgroup analyses in participants who did or did not experience falls in the past 12-months or had low or high physiological fall risk, concern about falling or executive function scores (median splits on PPA, Icon-FES and TMT-B) at baseline suggested no mediation on rate of falls (all p≥0.058; see Appendix 2). The subgroup analysis did suggest mediation by baseline status of physiological fall risk on physiological fall risk and by baseline status of concern about falling on concern about falling (p=0.004 and p=0.027 resp.; see Appendix 3). People with lower physiological fall risk at baseline had a significantly greater improvement of physiological fall risk at 6 months of 0.52 (95% CI 0.17-0.88) points. People with higher concern about falling at baseline had a significantly lower improvement of concern about falling at 12 months of -5 (95% CI -9, -1) points.

Process outcomes

Adverse events. Five falls occurred in three IG participants while exercising, which led to minor injuries (grazes, bruising, cuts). These falls were directly related to the intervention. Three falls occurred during exercise sessions and two were due to trips over exercise equipment.

Adherence. A total of 51 IG participants (20.1%) at 6-months, 81 (31.9%) at 12-months, 104 (40.9%) at 18-months, and 122 (48.0%) at 2-years had a median adherence of 0-minutes per week, either because of drop-out or non-usage attrition (see *Figure 1*). The remaining participants exercised for a median of 105.0 (IQR 58.5, n=203) minutes per week over the first 6-months, 114.0 (IQR 53.5, n=173) minutes per week over the first 12-months, 120.0 (IQR 39.3, n=150) minutes per week over 18-months and 120.3 (IQR 38.6, n=132) minutes per week over the full 2-years. Overall, 40.0%, 34.1%, 33.1% and 29.8% of IG participants achieved the prescribed dose over 6-, 12-, 18- and 24-months respectively.

Attitudes and usability. AFRIS and ESES scores at baseline were similar for both groups (p=0.595 and p=0.681 respectively) with medians of 42 (IQR 9) and 86 (IQR 22) in CG vs. 42 (IQR 8) and 87 (IQR 23) in IG. We repeated AFRIS and ESES and obtained PACES and SUS in IG at 6, 12, 18 and 24-months. AFRIS declined over time (p<0.0001) from a median of 42 (IQR 8) at baseline, to 40 (IQR 10) at 6-months, 39 (IQR 11) at 12-months, 39 (IQR 14) at 18-months and 35 (IQR 18) at 24-months, suggesting reduced intentions to continue the intervention. ESES also declined over time (p<0.0001) from a median of 87 (IQR 23) at baseline, to 75 (IQR 28) at 6-months, 70 (IQR 27) at 12-months, 69 (IQR 39) at 18-months and

59 (IQR 41) at 24-months, suggesting reduced exercise self-efficacy. PACES and SUS remained stable over time (p=0.362 and p=0.697 respectively) with medians of 27 (IQR 14) and 4.4 (IQR 0.8).

DISCUSSION

To our knowledge, this is the first large randomised controlled trial to evaluate the effects of a digital, unsupervised balance exercise programme on falls. We observed no significant effects on our primary outcomes, rate of falls and proportion of fallers, at 12-months. Yet, we did observe a significant reduction in fall rate of 16% and in proportion of injurious fallers of 20% at 24-months. Moreover, albeit not statistically significant (p=0.07), the effect size of a 18% reduction of fall rate at 12-months was similar to that at 24-months. The observed reduction in fall rate is comparable to that of previous studies, which achieved a 21% reduction with individually-delivered exercise programmes over 12-months and 14% over 24-months(2, 28). The 20% reduction in the proportion of injurious fallers at 24-months, on the other hand, may be higher than the previously reported 12% reductions(33). These findings indicate that technology can be used to deliver an e-Health balance exercise programme to older people which is effective at reducing fall rates and the proportion of injurious fallers over 24-months.

Secondary outcome analyses were not able to clearly highlight the pathway through which the reduction in falls rate and injurious fallers was achieved. In a subgroup of 226 participants, we observed a significant improvement in standing balance at 6 and 12-months; however, this was not confirmed through other balance and functional mobility measures. The *StandingTall* app includes a monthly balance assessment that comprises maintaining standing posture with feet in different positions, it is possible that the repeated practice carried over to laboratory assessments for IG participants. This trial might have been underpowered for detecting differences in fall risk factors, as our sample had a lower fall risk than anticipated. The a-priori sample size calculation was based on a sample with a mean physiological fall risk score (PPA) of 1.9 (SD 1.1) which is a full point higher than that of the current sample (PPA of 0.88, SD 0.88)(6). Interestingly, our pre-registered subgroup analyses found no significant modification of falls but did find indications of significant modification of the assessment outcomes, in people with lower physiological fall risk and lower concern about falling benefitting more. Quality of life measured with the EQ-5D-5L utility index also showed a significant improvement at 6-months, however no significant differences were found at 12 or 24-months.

Adherence to the intervention was good compared to previous exercise trials, with 40% of participants being fully adherent over the first 6-months and 30% being fully adherent over the full 2-years compared to pooled estimates of 21% in previous trials(31). Eighty percent of IG participants had a median adherence of 105-minutes over 6-months, and over half sustained a median adherence of 120-minutes over 24-months, despite the low level of contact during the study (two home visits in the first month and incidental follow-up calls to complete missing data). Adherence was collected automatically and is therefore a true representation of the actual dosage of balance training people received. Enjoyment and usability of the *StandingTall* intervention remained high throughout the entire study duration. Weekly medians suggest that the exercises might have become part of the lifestyle of those participants who remained in the study. While intentions and self-efficacy towards completing 2-hours of exercise per week declined over time, this is likely a more realistic reflection of actual long-term self-efficacy. The high adherence and zero serious adverse events support the feasibility and saefety of upscaling the intervention to a population level.

The strengths of this study were its large sample size, pragmatic design using a programme that could be delivered as part of routine care, broad inclusion criteria and use of methods designed to reduce the risk of bias such as concealed random allocation to groups, blinded outcome assessment, intention-to-treat analyses, and pre-registered statistical analysis plan (see PEDRO assessment in *Appendix 4*). The primary study limitations were the reliance on self-reported falls; however, the weekly e-diaries by both groups should have removed a reporting bias. Similar to many other exercise trials, participant masking was not possible. This might have led to bias by expectation, considering that many outcomes were self-rated. Thirdly, our study design intentionally included more than one outcome measure to account for the complex aetiology of falls, and, in theory, the subsequent multiple testing of the results could introduce error. Finally, it is possible that our weekly education fact sheets have induced a behaviour change in our control group, reducing our statistical power.

Novel methods for delivery of quality healthcare are required to increase effectiveness of fall prevention programs while containing costs and using scarce human resources to maximum effect. The ultimate success of a health promotion programme depends both on its effectiveness and its reach and acceptability in the community. A recently published multifactorial fall prevention trial in 5,451 older people at high risk of fall injuries illustrated that all participants had poor balance, and 95% agreed to take up an exercise program (30). Yet, the authors indicated that uptake and adherence to community-based exercise programs was low, and the evidence-base of these available exercise programs was uncertain (30). *StandingTall* fills an important gap by assisting older people to exercise at home, i.e. those who are unable (or unwilling) to attend out-of-house or group exercises, or those who wish to combine group and home-based exercises. Also, in light of the COVID-19 pandemic, as face-to-face delivery has been curtailed and de-conditioning is widespread, e-Health can offer an engaging, home-based substitute to reduce long-term adverse health consequences in older people from extended periods of isolation. E-Health programmes such as *StandingTall* can provide older people with an opportunity to stay active to prevent physical deconditioning and concomitant falls, functional dependence and increased healthcare use while maintaining COVID-19 safety recommendations.

In conclusion, our results show that a tailored e-Health exercise programme is an effective, low-resources, and thus low-cost, intervention towards the prevention of falls in older people. *StandingTall* is a scalable intervention and can be easily implemented into clinical practice, providing health professionals with a platform to remotely set-up, monitor and tailor the programme for their patients. *StandingTall* offers full user autonomy and requires minimal interaction with health professionals. An economic evaluation is planned to be undertaken to determine whether *StandingTall* represents value for money.

Contributor and guarantor information

Conceptualisation, KD, SRL, LC, GARZ, JCTC, BT; data collection, KD, TV, AW, JC, GM, LM, KSvS; data analysis, BT, NB, KSvS; data interpretation: KD, TL, KSvS; writing—original draft preparation, KD, KSvS; writing—review and editing, TV, SRL, LC, GARZ, JCTC, TL, AW, JC, GM, LM, BT, NB; final approval, KD, TV, SRL, LC, GARZ, JCTC, TL, AW, JC, GM, LM, BT, NB, KSvS, guarantor: KD. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Data sharing

Deidentified participant data may be accessed by researchers who provide a methodologically sound proposal. Proposals should be directed to k.delbaere@neura.edu.au and data are available from the date of publication of this manuscript. The study protocol is available as a free-access publication(6) and the statistical analysis plan is available on OpenScience framework (https://osf.io/42gje/).

Transparency statement

Kim Delbaere (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

Declaration of interests

The authors reported that the PPA (NeuRA FallScreen) is commercially available through Neuroscience Research Australia. No other conflicts of interest were reported.

Role of the funding source

This study was funded by Australian National Health and Medical Research Council grant APP1084739, Gandel Philaptropy and NeuRA Foundation. KD and SRL were supported by the Australian National Health and Medical Research Council (APP1105106, APP1117171). TL is currently supported by a NHMRC Early Career Fellowship (APP1141392) and National Heart Foundation Postdoctoral Fellowship (award ID 101956). KSvS was supported by a Human Frontier Science Program Fellowship (LT001080/2017). The funders did not have a role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

REFERENCES

1. James SL, Lucchesi LR, Bisignano C, Castle CD, Dingels ZV, Fox JT, et al. The global burden of falls: global, regional and national estimates of morbidity and mortality from the Global Burden of Disease Study 2017. Injury Prevention. 2020:injuryprev-2019-043286.
2. Sherrington C, Fairhall NJ, Wallbank GK, Tiedemann A, Michaleff ZA, Howard K, et al. Exercise for preventing falls in older people living in the community. Cochrane Database of Systematic Reviews. 2019(1).
3. Yardley L, Donovan-Hall M, Francis K, Todd C. Attitudes and beliefs that predict older people's intention to undertake strength and balance training. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2007;62(2):P119-P25.
4. Simek EM, McPhate L, Haines TP. Adherence to and efficacy of home exercise programs to prevent falls: A systematic review and meta-analysis of the impact of exercise program characteristics. Preventive Medicine. 2012;55(4):262-75.
5. Lamb SE, Jorstad-Stein EC, Hauer K, Becker C, Prevention of Falls Network E, Outcomes Consensus G. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. J Am Geriatr Soc. 2005;53(9):1618-22.
6. Delbaere K, Valenzuela T, Woodbury A, Davies T, Yeong J, Steffens D, et al. Evaluating the effectiveness of a home-based exercise programme delivered through a tablet computer for preventing falls in older community-dwelling people over 2 years: study protocol for the Standing Tall randomised controlled trial. BMJ Open. 2015;5(10):e009173.
7. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. Journal of the American Geriatrics Society. 1975;23(10):433-41.
8. Lord SR, Menz HB, Tiedemann A. A physiological profile approach to falls risk assessment and prevention. Physical therapy. 2003;83(3):237-52.
9. Csuka M, McCarty DJ. Simple method for measurement of lower extremity muscle strength. The American journal of medicine. 1985;78(1):77-81.
10. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. Journal of the American geriatrics Society. 1991;39(2):142-8.
11. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. Journal of gerontology. 1994;49(2):M85-M94.
12. Bohannon RW. Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants. Age and Ageing. 1997;26(1):15-9.
13. Schoene D, Smith ST, Davies TA, Delbaere K, Lord SR. A Stroop Stepping Test (SST) using low-cost computer game technology discriminates between older fallers and non-fallers. Age and Ageing. 2014;43(2):285-9.
14. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society. 2005;53(4):695-9.
15. Reitan RM, Wolfson D. The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation. Tucson, Ariz: Neuropsychology Press.; 1985.
16. Trenerry MR, Crosson B, DeBoe J, Leber W. Stroop neuropsychological screening test. Odessa, FL: Psychological Assessment Resources. 1989.
17. Delbaere K, T. Smith S, Lord SR. Development and Initial Validation of the Iconographical Falls Efficacy Scale. The Journals of Gerontology: Series A. 2011;66A(6):674-80.

18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. Journal of general internal medicine. 2001;16(9):606-13.
19. Gatt JM, Burton KL, Schofield PR, Bryant RA, Williams LM. The heritability of mental health and wellbeing defined using COMPAS-W, a new composite measure of wellbeing. Psychiatry Research. 2014;219(1):204-13.
20. Üstün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization disability assessment schedule 2.0. Bulletin of the World Health Organization. 2010;88:815-23.
21. Viney R, Norman R, King MT, Cronin P, Street DJ, Knox S, et al. Time trade-off derived EQ-5D weights for Australia. Value in Health. 2011;14(6):928-36.
22. Richardson JR, Peacock SJ, Hawthorne G, Iezzi A, Elsworth G, Day NA. Construction of the descriptive system for the assessment of quality of life AQoL-6D utility instrument. Health and quality of life outcomes. 2012;10(1):38.
23. Delbaere K, Hauer K, Lord SR. Evaluation of the incidental and planned activity questionnaire for older people. British journal of sports medicine. 2010;44(14):1029-34.
24. van Schooten KS, Rispens SM, Elders PJ, Lips P, van Dieen JH, Pijnappels M. Assessing physical activity in older adults: required days of trunk accelerometer measurements for reliable estimation. Journal of aging and physical activity. 2015;23(1):9-17.
25. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies. 2009;4(3):114-23.
26. Kendzierski D, DeCarlo KJ. Physical activity enjoyment scale: Two validation studies. Journal of sport & exercise psychology. 1991;13(1).
27. Everett B, Salamonson Y, Davidson PM. Bandura's exercise self-efficacy scale: validation in an Australian cardiac rehabilitation setting. Int J Nurs Stud. 2009;46(6):824-9.
28. Finnegan S, Seers K, Bruce J. Long-term follow-up of exercise interventions aimed at preventing falls in older people living in the community: a systematic review and meta-analysis. Physiotherapy. 2019;105(2):187-99.
29. Zhao R, Bu W, Chen X. The efficacy and safety of exercise for prevention of fall-related injuries in older people with different health conditions, and differing intervention protocols: a meta-analysis of randomized controlled trials. BMC Geriatrics. 2019;19(1):341.
30. Bhasin S, Gill TM, Reuben DB, Latham NK, Ganz DA, Greene EJ, et al. A Randomized Trial of a Multifactorial Strategy to Prevent Serious Fall Injuries. New England Journal of Medicine. 2020;383(2):129-40.

FIGURES AND TABLES

Figure 1: Flowchart of study recruitment and retention.

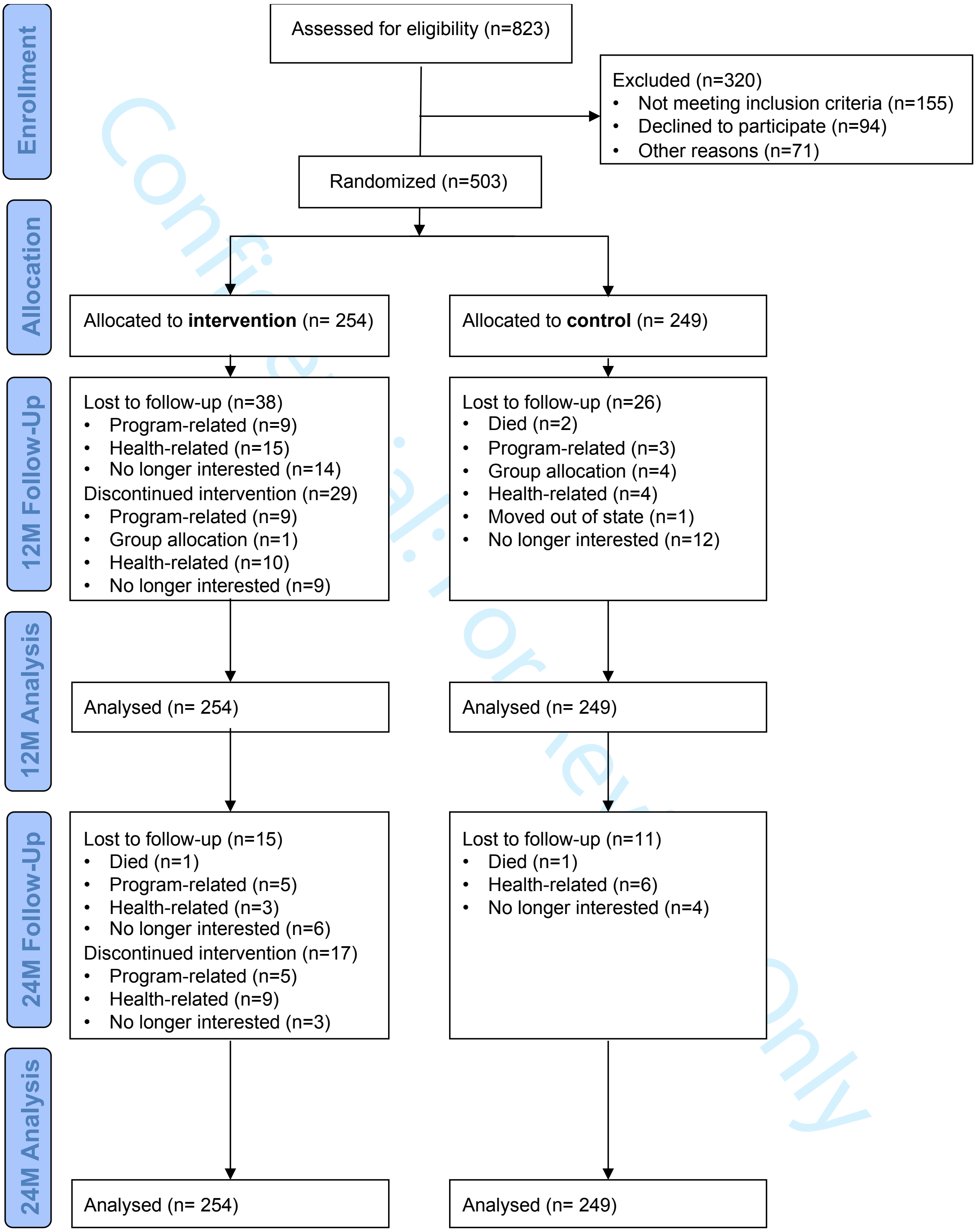

Figure 2: Effect on rate of falls and faller status. Primary outcomes are bolded, values indicate incidence rate ratio (IRR) or relative risk (RR) with corresponding 95% confidence interval (95% CI). Vertical line indicates no difference between the groups (i.e. IRR or RR of 1). CACE shows complier average causal effect.

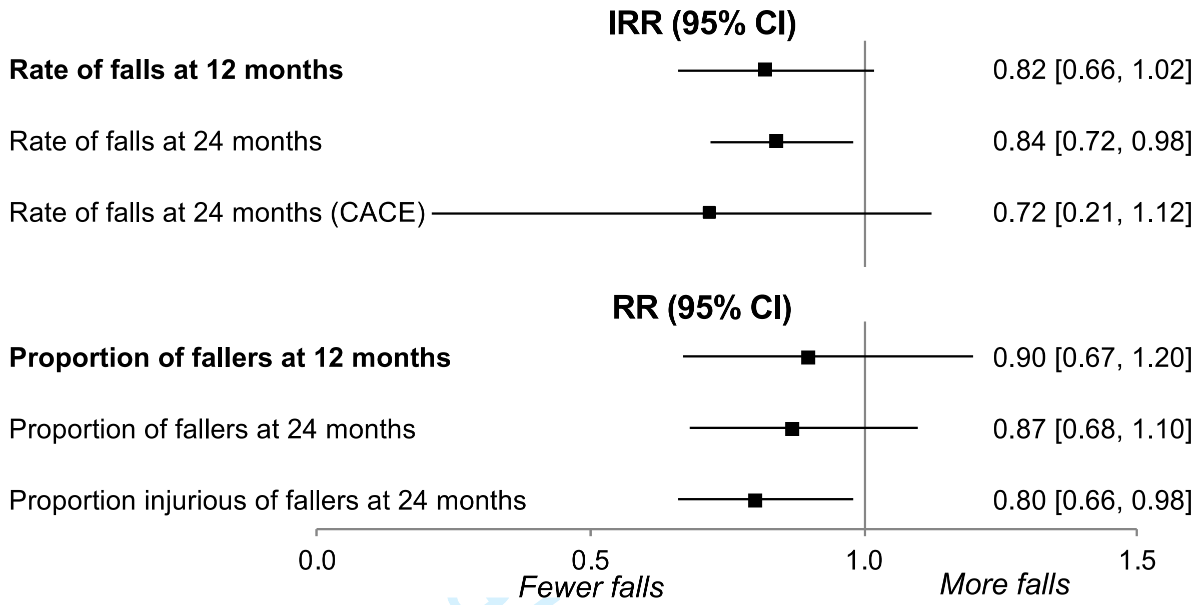

Table 1: Baseline characteristics of all participants (N=503)

Variable	Intervention (n=254)	Control (n=249)
Age (years)	77.1 (5.5)	77.7 (5.5)
Female gender (%)	177 (69.7%)	162 (65.1%)
BMI (kg/m^2)	27.3 (4.5)	27.0 (4.9)
Education (years)	14.4 (4.1)	14.6 (4.4)
Living alone (%)	113 (44.5%)	104 (41.9%)
Owns a computer (%)	214 (85.0%)	220 (88.4%)
Uses walking aid (%)	18 (7.1%)	20 (8.0%)
Falls in previous year (number)	0 [1]	0 [1]
EQ-5D-5L VAS (score range 0 - 100)	90 [15]	85 [15]
Medical conditions (number)	0 [1]	0 [1]
Prescription medication (number)#	3 [3]	3 [3]
Montreal Cognitive Assessment (score range 0 - 30)	27 [3]	27 [3]
Trail Making Test B minus A (TMT-B min TMT-A; seconds)	55.3 [36.2]	54.8 [44.8]
Patient Health Questionnaire (PHQ-9; score range 0 - 27)	2 [4]	2 [4]
Iconographical Falls Efficacy Scale (icon-FES; score range 30 - 120)	53 (16)	55 (16)
Physiological fall risk (PPA score)	0.99 (0.74)	1.19 (0.87)
Timed up and go (seconds)	8.5 (3.3)	8.6 (3.0)

Note: values are mean (standard deviation), absolute (relative %) or median [IQR]; end of range indicating best score is underlined.
#available for 335 (66.6%) of people.

Table 2: Effect on secondary outcome measures in all participants

Variable better scores underlined	Control group Mean (SD)					Intervention group Mean (SD)					Change in IG compared to CG Beta (95% confidence interval), p			
	0M	**6M**	**12M**	**18M**	**24M**	**0M**	**6M**	**12M**	**18M**	**24M**	**0-6M**	**0-12M**	**0-18M**	**0-24M**
Psychological wellbeing														
PHQ-9 (score range 0 - 27)	2 [4]	3 [5]	3 [4]	4 [3]	3 [5]	2 [4]	2 [4]	3 [4]	3 [5]	3 [4]	0 (0, 0), p=NA	0 (0, 0), p=NA	0 (0, 0), p=NA	0 (0, 0), p=NA
Icon-FES (score range 30 - 120)	55 (16)	53 (16)	55 (16)	57 (20)	58 (18)	53 (16)	52 (17)	51 (16)	52 (19)	53 (18)	2 (-2, 6) p=0.331	-1 (-5, 3) p=0.597	-2 (-6, 2) p=0.254	-2 (-6, 2) p=0.400
COMPAS-W (score range 26 - 130)	100 (11)	100 (11)	100 (11)	100 (12)	100 (11)	101 (12)	102 (12)	102 (12)	102 (13)	102 (12)	1 (-1, 4) p=0.342	1 (-1, 4) p=0.324	1 (-2, 3) p=0.614	1 (-1, 4) p=0.313
Health-related quality of life														
WHODAS (score range 0 – 100%)	6.3 [12.5]	7.2 [15.6]	8.8 [14.7]	8.1 [16.5]	6.3 [16.5]	4.1 [10.4]	4.9 [12.9]	6.3 [13.4]	6.2 [14.5]	7.7 [14.6]	-1.6 (-4.1, 0.8), p=NA	-0.2 (-2.7, 2.3), p=NA	0.3 (-2.5, 3.1), p=NA	0.2 (-2.0, 2.5), p=NA
EQ-5D-5L VAS (score range 0 - 100)	85 [15]	87 [16]	83 [19]	80 [16]	80 [20]	90 [15]	90 [14]	89 [11]	87 [15]	88 [14]	-2 (-6, 3), p=NA	0 (-4, 4), p=NA	-2 (-6, 2), p=NA	1 (-4, 6), p=NA
EQ-5D-5L utility (score range 0 - 1)	0.89 [0.03]	0.87 [0.04]	0.87 [0.05]	0.86 [0.06]	0.86 [0.05]	0.89 [0.04]	0.87 [0.05]	0.87 [0.05]	0.87 [0.06]	0.88 [0.05]	**0.03 (0.01, 0.06), p=NA**	-0.01 (-0.04, 0.03), p=NA	0.01 (-0.02, 0.05), p=NA	0.01 (-0.02, 0.04), p=NA
AQOL-6D (utility score range 0 - 1)	0.88 [0.16]	0.87 [0.15]	0.87 [0.16]	0.86 [0.20]	0.86 [0.16]	0.89 [0.14]	0.90 [0.15]	0.89 [0.16]	0.90 [0.19]	0.89 [0.15]	0.01 (-0.02, 0.04), p=NA	-0.01 (-0.04, 0.02), p=NA	0.14 (-0.94, 1.24), p=NA	0.01 (-0.02, 0.03), p=NA
Physical activity levels														
IPEQ planned activity (hrs)	5.3 [6.7]	5.0 [7.6]	4.9 [5.8]	4.2 [7.3]	4.4 [6.0]	5.4 [7.6]	5.7 [9.1]	5.9 [7.4]	5.0 [7.4]	5.0 [6.5]	0.5 (-0.8, 1.8), p=NA	0 (-4.3, 4.2), p=NA	-1.6 (-5.6, 2.4), p=NA	1.1 (-3.6, 5.7), p=NA
IPEQ incidental activity (hrs)	34.0 [28.9]	38.0 [30.1]	32.0 [33.6]	36.9 [24.3]	33.8 [35.4]	33.0 [35.2]	38.0 [32.9]	32.0 [33.8]	36.9 [24.4]	37.8 [34.6]	-2.8 (-7.8, 2.3), p=NA	-0.9 (-6.8, 5.0), p=NA	-1.7 (-7.0, 3.7), p=NA	-1.9 (-7.3, 3.4), p=NA

IPEQ planned exercise (hrs)	3.0 [4.9]	3.1 [4.7]	3.0 [4.5]	2.7 [4.5]	2.6 [3.9]	2.8 [4.9]	3.5 [5.3]	3.6 [4.9]	2.7 [4.4]	3.0 [4.7]	0.6 (-0.4, 1.7), p=NA	0.9 (-0.1, 2.0), p=NA	0.1 (-0.9, 1.2), p=NA	0.6 (-0.5-1.6), p=NA
MM Walking time (hrs)	1.26 (0.47)	1.20 (0.47)	1.16 (0.51)	1.13 (0.49)	1.19 (0.54)	1.35 (0.59)	1.31 (0.54)	1.25 (0.61)	1.21 (0.55)	1.21 (0.61)	0.02 (-0.11, 0.14) p=0.790	0.01 (-0.13, 0.14) p=0.901	-0.01 (-0.14, 0.12) p=0.875	-0.07 (-0.19, 0.06) p=0.301
MM Walking bouts	423 (142)	429 (153)	418 (163)	396 (161)	411 (175)	441 (162)	459 (182)	444 (179)	424 (171)	413 (184)	11 (-31, 54) p=0.583	8 (-34, 51) p=0.692	10 (-36, 56) p=0.644	-16 (-58, 27) p=0.452
MM Standing time (hrs)	2.52 (0.75)	2.55 (0.92)	2.52 (0.87)	2.42 (0.86)	2.44 (0.93)	2.59 (0.87)	2.65 (1.11)	2.62 (0.99)	2.48 (1.01)	2.34 (0.96)	0.03 (-0.23, 0.30) p=0.788	0.03 (-0.23, 0.28) p=0.825	-0.01 (-0.26, 0.23) p=0.913	-0.16 (-0.41, 0.09) p=0.190
MM Standing bouts	870 (289)	880 (314)	877 (312)	824 (297)	848 (353)	880 (320)	948 (366)	934 (361)	876 (355)	825 (355)	59 (-24, 142) p=0.156	47 (-38, 132) p=0.263	42 (-48, 132) p=0.338	-32 (-116, 52) p=0.433

Note: values are mean (standard deviation) or median [IQR]; end of range indicating best score is underlined; p=NA indicates bootstrapped outcomes, which did not allow us to estimate p-values.

PHQ-9: nine-item Patient Health Questionnaire; Icon-FES: Iconographical Falls Efficacy Scale; COMPAS-W: COMPAS-W scale; WHODAS: 12-item WHO Disability Assessment Schedule; EQ-5D-5L: 5-level EuroQol- 5 Dimension; AQOL-6D: 20-item Assessment of Quality of Life 6-Dimensions; IPEQ: Incidental and Planned Exercise Questionnaire; MM: McRoberts MoveMonitor.

Table 3: Effect on secondary outcome measures in a subsample of 226 participants

Variable	Control group Mean (SD)			Intervention group Mean (SD)			Change in intervention compared to control group Beta (95% CI), p	
	0M	**6M**	**12M**	**0M**	**6M**	**12M**	**0-6M**	**0-12M**
Physiological fall risk								
PPA score	1.19 (0.87)	1.17 (0.77)	0.97 (0.93)	0.99 (0.74)	0.82 (0.82)	0.76 (0.92)	-0.15 (-0.39, 0.09) p=0.214	-0.01 (-0.27, 0.26) p=0.955
Balance, functional mobility and gait								
Standing balance (s)	188 [69]	189 [81]	186 [73]	193 [95]	209 [71]	198 [70]	**11 (3, 19),** p=NA	**10 (1, 19),** p=NA
Maximum lean range AP (cm)	15 (3)	16 (4)	18 (4)	15 (3)	17 (4)	19 (4)	1 (0, 2) p=0.206	1 (0, 2) p=0.213
Coordinated lean (score)	7 [13]	9 [15]	8 [11]	7 [11]	6 [11]	5 [12]	-2 (-4, 0), p=NA	-1 (-3, 1), p=NA
Timed up and go (s)	8.6 (3.0)	8.7 (3.6)	8.6 (4.1)	8.5 (3.3)	8.5 (3.4)	8.2 (3.3)	-1.8 (-4.4, 0.7) p=0.146	-1.5 (-3.7, 0.8) p=0.190
5-times sit-to-stand (s)	12.4 (4.3)	12.1 (5.2)	11.5 (4.7)	12.4 (5.4)	12.1 (4.6)	11.0 (3.8)	0.1 (-1.0, 1.2) p=0.864	-0.4 (-1.5, 0.6) p=0.411
10-m walk (s)	9.0 (2.0)	9.1 (3.1)	8.8 (3.0)	8.9 (2.1)	8.7 (2.4)	8.6 (2.5)	-0.3 (-0.8, 0.3) p=0.322	-0.1 (-0.6, 0.5) p=0.802
Short Physical Performance Battery (score)	11 [2]	11 [2]	11 [2]	11 [2]	11 [1]	11 [2]	0 (0, 0), p=NA	0 (0, 1), p=NA
Stepping performance								
Choice stepping reaction time (s)	1.16 (0.20)	1.17 (0.17)	1.18 (0.23)	1.13 (0.18)	1.15 (0.19)	1.17 (0.18)	0.01 (-0.05, 0.06) p=0.744	0.03 (-0.03, 0.08) p=0.380
Inhibitory stepping reaction time (s)	1.32 (0.40)	1.32 (0.43)	1.32 (0.38)	1.26 (0.37)	1.36 (0.49)	1.29 (0.36)	0.10 (-0.04, 0.23), p=0.143	0.03 (-0.11, 0.17), p=0.645
Stroop stepping reaction time (s)	1.25 (0.42)	1.24 (0.39)	1.22 (0.39)	1.21 (0.34)	1.28 (0.38)	1.19 (0.34)	0.17 (0.16) p=0.302	0.26 (0.07) p=0.116
Cognitive performance and executive functions								
TMT-A (s)	31.9 [11.8]	27.5 [14.4]	39.6 [12.8]	29.8 [12.7]	29.7 [11.8]	28.4 [13.9]	1.8 (-0.7, 4.2), p=NA	0.6 (-2.0, 3.1), p=NA
TMT-B (s)	85.2 [50.7]	90.1 [54.3]	84.3 [55.8]	87.7 [43.2]	87.5 [53.5]	87.8 [51.8]	1.5 (-8.0, 11.0), p=NA	5.1 (-5.1, 15.3), p=NA
TMT-B min TMT-A (s)	54.8 [44.8]	60.5 [44.0]	56.2 [44.1]	55.3 [36.2]	55.9 [38.9]	59.7 [41.9]	-0.3 (-10.5, 9.6), p=NA	4.5 (-5.8, 15.0), p=NA
Victoria Stroop ratio	2.13 (0.87)	1.87 (0.92)	1.98 (0.91)	1.95 (0.74)	1.89 (0.89)	1.98 (0.96)	0.20 (-0.13, 0.53) p=0.224	0.18 (-0.15, 0.51) p=0.270
Victoria Stroop errors	4 [5]	3 [5]	3 [4]	3 [5]	3 [5]	2 [4]	0 (0, 0) p=NA	0 (0, 0) p=NA

Note: values are mean (standard deviation) or median [IQR]; p=NA indicates bootstrapped outcomes, which did not allow us to estimate p-values.
PPA: Physiological Profile Assessment; AP: anteroposterior; TMT: Trail-Making Test.

ONLINE SUPPLEMENTARY MATERIAL

Online Appendix 1: **Negative binominal vs Poisson results**

	Poisson			**Negative binominal**		
	AIC	*IRR (95% CI)*	*p*	*AIC*	*IRR (95% CI)*	*p*
Fall rate at 12-months	1220.50	0.84 (0.62, 1.13)	0.0710	1123.61	0.82 (0.66, 1.02)	0.2526
Fall rate at 24-months	1808.06	0.84 (0.72, 0.98)	0.0273	1530.52	0.86 (0.67, 1.11)	0.2516

Online Appendix 2: Fall incidence rate ratios per pre-specified subgroups with statistical testing of differences.

12-months history of falls	**No past falls n=307**	**Past falls n=192**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.78 (0.56, 1.09), *p=0.142*	0.77 (0.58, 1.01), *p=0.063*	0.931
Rate of falls at 24-months	0.91 (0.72, 1.15), *p*=0.438	0.67 (0.55, 0.83), *p<0.001*	0.058

Physiological fall risk	**PPA ≤0.823 points n=245**	**PPA >0.823 points n=258**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.81 (0.58, 1.14), *p=0.227*	0.88 (0.66, 1.16), *p=0.348*	0.163
Rate of falls at 24-months	0.74 (0.58, 0.94), *p=0.015*	0.92 (0.76, 1.13), *p=0.448*	0.740

Executive function	**TMT-B ≤86.4 s n=263**	**TMT-B >86.4 s n=240**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.93 (0.69, 1.25), *p=0.637*	0.75 (0.55, 1.03), *p=0.074*	0.718
Rate of falls at 24-months	0.86 (0.69, 1.06), *p=0.166*	0.81 (0.65, 1.02), *p=0.072*	0.331

Concern about falling	**iconFES ≤49 n=244**	**iconFES >49 n=259**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.95 (0.68, 1.32), *p=0.750*	0.80 (0.60, 1.06), *p= 0.120*	0.447
Rate of falls at 24-months	0.91 (0.72, 1.15), *p=0.422*	0.80 (0.65, 0.99), *p=0.042*	0.438

Note: PPA: Physiological profile assessment, TMT-B: Trail Making Test part B, iconFES: Iconographical Fall Efficacy Scale. Cutpoints for PPA, TMT-B and iconFES are based on a median split.

Online Appendix 3: Change in subgroups (N=226) over first 12-months

Variable	Change in intervention compared to control group Beta (95% CI), p		Change in intervention compared to control group Beta (95% CI), p		Difference in change between subgroups Beta (95% CI), p	
	0-6M	0-12M	0-6M	0-12M	0-6M	0-12M
12-month history of falls	**No past falls n=91**		**Past falls n=131**			
PPA score	-0.24 (-0.50, 0.02), p=0.076	-0.05 (-0.32, 0.21), p=0.690	-0.32 (-0.65, 0.01), p=0.057	-0.26 (-0.59, 0.08), p=0.135	-0.08 (-0.50, 0.34), p=0.704	-0.20 (-0.63, 0.22), p=0.354
TMT-B (s)	-1.7 (-13.5, 9.5), p=NA	5.5 (-6.8, 18.0), p=NA	12.5 (-6.6, 31.9), p=NA	11.2 (-8.6, 32.2), p=NA	14.2 (-7.5, 36.6), p=NA	5.7 (-18.3, 30.3), p=NA
IconFES	-2 (-4, 0), p=0.074	0 (-2, 3), p=0.754	0 (-3, 4), p=0.964	-2 (-6, 1), p=0.226	0 (-5, 4), p=0.866	-3 (-7, 2), p=0.247
Physio-logical fall risk	**PPA ≤1.045 points n=113**		**PPA >1.045 points n=113**			
PPA score	**-0.64 (-0.91, -0.36), p<0.0001**	-0.37 (-0.64, -0.09), p=0.010	-0.11 (-0.34, 0.12), p=0.335	-0.04 (0.12), p=0.747	**0.52 (0.17, 0.88), p=0.004**	0.33 (0.69, -0.03), p=0.077
TMT-B (s)	-1.67 (-13.29, 10.47), p=NA	3.09 (-15.01, 19.72), p=NA	7.26 (-8.59, 22.53), p=NA	9.71 (-3.69, 23.90), p=NA	8.93 (-10.97, 28.06), p=NA	6.62 (-15.76, 27.70), p=NA
iconFES	-1 (-3, 1), p=0.488	-2 (-5, 1), p=0.272	2 (-1, 5), p=0.144	1 (-2, 4), p=0.510	4 (1, -8), p=0.097	3 (2, 7), p=0.217
Executive function	**TMT-B >86.85 s n=113**		**TMT-B ≤86.85 s n=113**			
PPA score	-0.14 (-0.42, 0.15), p=0.345	-0.03 (-0.31, 0.27), p=0.859	-0.42 (-0.71, -0.13), p=0.004	-0.24 (-0.53, 0.05) p=0.103	-0.28 (-0.69, 0.12), p=0.174	-0.21 (-0.62, 0.20), p=0.311
TMT-B (s)	12.3 (-5.7, 28.6), p=NA	16.4 (-3.0, 35.5), p=NA	-3.8 (-12.6, 5.0), p=NA	0.61 (-9,9, 10.2), p=NA	-16.0 (-35.6, 4.3), p=NA	-15.8 (-36.7, -5.8), p=NA
iconFES	2 (-1, 5), p=0.262	0 (-3, 4), p=0.839	-1 (-3, 1), p=0.332	-1 (-4, 2), p=0.474	-3 (-7, 2), p=0.205	-1 (-6, 3), p=0.521
Concern about falling	**iconFES >50 n=113**		**iconFES ≤50 n=113**			
PPA score	-0.36 (-0.64, -0.07) p=0.014	-0.13 (-0.41, 0.16, 0.16) p=0.389	-0.19 (-0.48, 0.11) p=0.212	-0.14 (0.15) p=0.379	0.17 (-0.24, 0.58), p=0.420	-0.01 (-0.42, 0.40), p=0.967
TMT-B (s)	10.2 (-6.1, 26.8), p=NA	17.4 (0.7, 35.1), p=NA	-5.0 (-16.5, 5.9), p=NA	-3.1 (-15.6, 8.4), p=NA	-15.2 (-34.9, 3.9), p=NA	-20.5 (-41.8, 0.5), p=NA
iconFES	1 (-2, 4), p=0.356	2 (-1, 5), p=0.213	3 (0, 5), p=0.019	-3 (-6, 0), p=0.058	2 (2, 7), p=0.265	**-5 (-9, -1), p=0.027**

Note: PPA: Physiological profile assessment, TMT-B: Trail Making Test part B, iconFES: Iconographical Fall Efficacy Scale. Cutpoints for PPA, TMT-B and iconFES are based on a median split. p=NA indicates bootstrapped outcomes, which did not allow us to estimate p-values.

Online Appendix 4: PEDro assessment

Item		**Response**	**Score**
1	Eligibility criteria were specified	Page 6	1
2	Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received)	Page 6	1
3	Allocation was concealed	Page 6	1
4	The groups were similar at baseline regarding the most important prognostic indicators	Table 1	1
5	There was blinding of all subjects	No, not possible in an exercise vs. control program	0
6	There was blinding of all therapists who administered the therapy	N/A, the programme was unsupervised	1
7	There was blinding of all assessors who measured at least one key outcome	Page 6	1
8	Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups	Page 9	1
9	All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by "intention to treat"	Page 8	1
10	The results of between-group statistical comparisons are reported for at least one key outcome	Figure 2, Table 2 & 3	1
11	The study provides both point measures and measures of variability for at least one key outcome	Figure 2, Table 2 & 3	1
Total Score			**10**

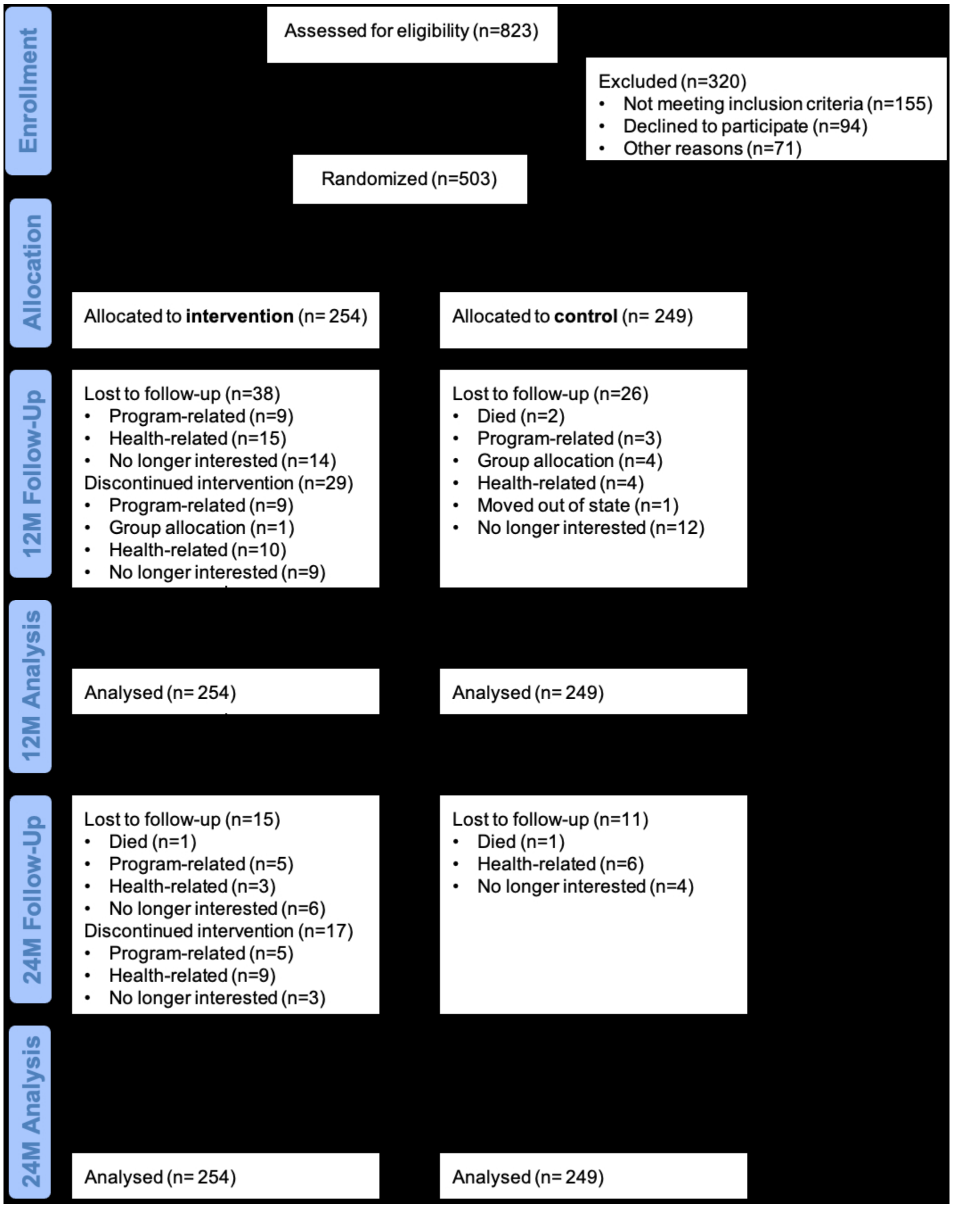

Figure 1: Flowchart of study recruitment and retention.

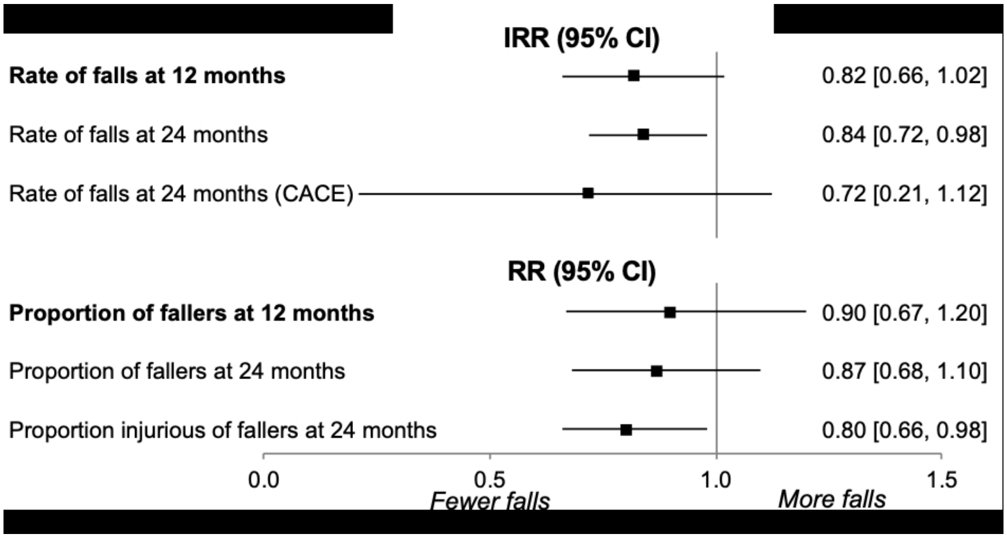

Figure 2: Effect on rate of falls and faller status. Primary outcomes are bolded, values indicate incidence rate ratio (IRR) or relative risk (RR) with corresponding 95% confidence interval (95% CI). Vertical line indicates no difference between the groups (i.e. IRR or RR of 1). CACE shows complier average causal effect.

ONLINE SUPPLEMENTARY MATERIAL

Online Appendix 1: **Negative binominal vs Poisson results**

	Poisson			Negative binominal		
	AIC	*IRR (95% CI)*	*p*	*AIC*	*IRR (95% CI)*	*p*
Fall rate at 12-months	1220.50	0.84 (0.62, 1.13)	0.0710	1123.61	0.82 (0.66, 1.02)	0.2526
Fall rate at 24-months	1808.06	0.84 (0.72, 0.98)	0.0273	1530.52	0.86 (0.67, 1.11)	0.2516

Online Appendix 2: Fall incidence rate ratios per pre-specified subgroups with statistical testing of differences.

12-months history of falls	**No past falls n=307**	**Past falls n=192**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.78 (0.56, 1.09), *p=0.142*	0.77 (0.58, 1.01), *p=0.063*	0.931
Rate of falls at 24-months	0.91 (0.72, 1.15), *p*=0.438	0.67 (0.55, 0.83), *p<0.001*	0.058

Physiological fall risk	**PPA ≤0.823 points n=245**	**PPA >0.823 points n=258**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.81 (0.58, 1.14), *p=0.227*	0.88 (0.66, 1.16), *p=0.348*	0.163
Rate of falls at 24-months	0.74 (0.58, 0.94), *p=0.015*	0.92 (0.76, 1.13), *p=0.448*	0.740

Executive function	**TMT-B ≤86.4 s n=263**	**TMT-B >86.4 s n=240**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.93 (0.69, 1.25), *p=0.637*	0.75 (0.55, 1.03), *p=0.074*	0.718
Rate of falls at 24-months	0.86 (0.69, 1.06), *p=0.166*	0.81 (0.65, 1.02), *p=0.072*	0.331

Concern about falling	**iconFES ≤49 n=244**	**iconFES >49 n=259**	**Interaction (*p*-value)**
Rate of falls at 12-months	0.95 (0.68, 1.32), *p=0.750*	0.80 (0.60, 1.06), *p= 0.120*	0.447
Rate of falls at 24-months	0.91 (0.72, 1.15), *p=0.422*	0.80 (0.65, 0.99), *p=0.042*	0.438

Note: PPA: Physiological profile assessment, TMT-B: Trail Making Test part B, iconFES: Iconographical Fall Efficacy Scale. Cutpoints for PPA, TMT-B and iconFES are based on a median split.

Online Appendix 3: Change in subgroups (N=226) over first 12-months

Variable	**Change in intervention compared to control group** Beta (95% CI), p		**Change in intervention compared to control group** Beta (95% CI), p		**Difference in change between subgroups** Beta (95% CI), p	
	0-6M	0-12M	0-6M	0-12M	0-6M	0-12M
12-month history of falls	**No past falls n=91**		**Past falls n=131**			
PPA score	-0.24 (-0.50, 0.02), p=0.076	-0.05 (-0.32, 0.21), p=0.690	-0.32 (-0.65, 0.01), p=0.057	-0.26 (-0.59, 0.08), p=0.135	-0.08 (-0.50, 0.34), p=0.704	-0.20 (-0.63, 0.22), p=0.354
TMT-B (s)	-1.7 (-13.5, 9.5), p=NA	5.5 (-6.8, 18.0), p=NA	12.5 (-6.6, 31.9), p=NA	11.2 (-8.6, 32.2), p=NA	14.2 (-7.5, 36.6), p=NA	5.7 (-18.3, 30.3), p=NA
IconFES	-2 (-4, 0), p=0.074	0 (-2, 3), p=0.754	0 (-3, 4), p=0.964	-2 (-6, 1), p=0.226	0 (-5, 4), p=0.866	-3 (-7, 2), p=0.247
Physio-logical fall risk	**PPA ≤1.045 points n=113**		**PPA >1.045 points n=113**			
PPA score	**-0.64 (-0.91, -0.36), p<0.0001**	-0.37 (-0.64, -0.09), p=0.010	-0.11 (-0.34, 0.12), p=0.335	-0.04 (0.12), p=0.747	**0.52 (0.17, 0.88), p=0.004**	0.33 (0.69, -0.03), p=0.077
TMT-B (s)	-1.67 (-13.29, 10.47), p=NA	3.09 (-15.01, 19.72), p=NA	7.26 (-8.59, 22.53), p=NA	9.71 (-3.69, 23.90), p=NA	8.93 (-10.97, 28.06), p=NA	6.62 (-15.76, 27.70), p=NA
iconFES	-1 (-3, 1), p=0.488	-2 (-5, 1), p=0.272	2 (-1, 5), p=0.144	1 (-2, 4), p=0.510	4 (1, -8), p=0.097	3 (2, 7), p=0.217
Executive function	**TMT-B >86.85 s n=113**		**TMT-B ≤86.85 s n=113**			
PPA score	-0.14 (-0.42, 0.15), p=0.345	-0.03 (-0.31, 0.27), p=0.859	-0.42 (-0.71, -0.13), p=0.004	-0.24 (-0.53, 0.05) p=0.103	-0.28 (-0.69, 0.12), p=0.174	-0.21 (-0.62, 0.20), p=0.311
TMT-B (s)	12.3 (-5.7, 28.6), p=NA	16.4 (-3.0, 35.5), p=NA	-3.8 (-12.6, 5.0), p=NA	0.61 (-9,9, 10.2), p=NA	-16.0 (-35.6, 4.3), p=NA	-15.8 (-36.7, -5.8), p=NA
iconFES	2 (-1, 5), p=0.262	0 (-3, 4), p=0.839	-1 (-3, 1), p=0.332	-1 (-4, 2), p=0.474	-3 (-7, 2), p=0.205	-1 (-6, 3), p=0.521
Concern about falling	**iconFES >50 n=113**		**iconFES ≤50 n=113**			
PPA score	-0.36 (-0.64, -0.07) p=0.014	-0.13 (-0.41, 0.16, 0.16) p=0.389	-0.19 (-0.48, 0.11) p=0.212	-0.14 (0.15) p=0.379	0.17 (-0.24, 0.58), p=0.420	-0.01 (-0.42, 0.40), p=0.967
TMT-B (s)	10.2 (-6.1, 26.8), p=NA	17.4 (0.7, 35.1), p=NA	-5.0 (-16.5, 5.9), p=NA	-3.1 (-15.6, 8.4), p=NA	-15.2 (-34.9, 3.9), p=NA	-20.5 (-41.8, 0.5), p=NA
iconFES	1 (-2, 4), p=0.356	2 (-1, 5), p=0.213	3 (0, 5), p=0.019	-3 (-6, 0), p=0.058	2 (2, 7), p=0.265	**-5 (-9, -1), p=0.027**

Note: PPA: Physiological profile assessment, TMT-B: Trail Making Test part B, iconFES: Iconographical Fall Efficacy Scale. Cutpoints for PPA, TMT-B and iconFES are based on a median split. p=NA indicates bootstrapped outcomes, which did not allow us to estimate p-values.

Online Appendix 4: PEDro assessment

Item		Response	Score
1	Eligibility criteria were specified	Page 6	1
2	Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received)	Page 6	1
3	Allocation was concealed	Page 6	1
4	The groups were similar at baseline regarding the most important prognostic indicators	Table 1	1
5	There was blinding of all subjects	No, not possible in an exercise vs. control program	0
6	There was blinding of all therapists who administered the therapy	N/A, the programme was unsupervised	1
7	There was blinding of all assessors who measured at least one key outcome	Page 6	1
8	Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups	Page 9	1
9	All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by "intention to treat"	Page 8	1
10	The results of between-group statistical comparisons are reported for at least one key outcome	Figure 2, Table 2 & 3	1
11	The study provides both point measures and measures of variability for at least one key outcome	Figure 2, Table 2 & 3	1
Total Score			**10**