

**Both baseline and change in lower limb muscle strength in younger women are independent predictors of balance in middle-age: a 12-yr population-based prospective study**

Feitong Wu<sup>1</sup>, Michele Callisaya<sup>1</sup>, Karen Wills<sup>1</sup>, Laura L Laslett<sup>1</sup>, Graeme Jones<sup>1</sup>, Tania Winzenberg<sup>1,2</sup>

<sup>1</sup>Menzies Institute for Medical Research, University of Tasmania, Private Bag 23, Hobart, 7000 Tasmania, Australia

<sup>2</sup>Faculty of Health, University of Tasmania, Hobart, Tasmania, Australia

**Email**

Feitong.Wu@utas.edu.au

Michele.Callisaya@utas.edu.au

Karen.Wills@utas.edu.au

Laura.Laslett@utas.edu.au

Graeme.Jones@utas.edu.au

Tania.Winzenberg@utas.edu.au

**\*Corresponding Author:**

Tania Winzenberg, PhD

Menzies Institute for Medical Research University of Tasmania

17 Liverpool Street, Hobart Tasmania 7000

+61(0)362267770

[Tania.Winzenberg@utas.edu.au](mailto:Tania.Winzenberg@utas.edu.au)

**No supplemental data included.**

**Disclosure:** All authors state that they have no conflicts of interest.

## Abstract

Poor balance is a risk factor for falls and fracture in older adults, but little is known about modifiable factors affecting balance in younger women. This study aimed to examine whether lower limb muscle strength (LMS) in young women and changes in LMS are independent predictors of balance in middle-age.

This was an observational 10-yr follow-up of 470 women aged 25-44 years at baseline who had previously participated in a 2-yr population-based randomised controlled trial of osteoporosis education interventions. Linear regression was used to examine the association between baseline LMS (by dynamometer) and change in LMS over 12 years with balance at 12 years (timed up and go test (TUG), step test (ST), functional reach test (FRT) and lateral reach test (LRT)).

LMS declined by a mean of 17.3 kg over 12 years. After adjustment for potential confounders, baseline and change in LMS were independently beneficially associated with TUG ( $\beta=-0.008\text{sec/kg}$ , 95% confidence interval (CI): -0.01 to -0.006 and -0.006sec/kg: -0.009 to -0.003 for baseline and change respectively), FRT ( $\beta=0.057\text{cm/kg}$ , 95%CI: 0.030 to 0.084 and 0.071cm/kg: 0.042 to 0.101) and LRT ( $\beta=0.030\text{cm/kg}$ , 95%CI: 0.012 to 0.049 and 0.022cm/kg: 0.002 to 0.043) 12 years later. There was an association between baseline LMS and ST ( $\beta=0.044\text{steps/kg}$ , 95%CI: 0.022 to 0.067) but not between change in LMS and ST.

Among young women, greater LMS at baseline and slower decline over time are both associated with better balance in midlife. Analogous to the contributions of peak bone mass and bone loss to fracture risk in older adults, this suggests that both improvement of muscle strength in younger age and prevention of age-related loss of muscle strength could be potentially useful strategies to improve balance and reduce falls in later life.

**KEY WORDS:** Muscle strength, Muscle loss, Balance, Younger women, Longitudinal

## Introduction

Falls are a major health issue among older adults, with about one in three community-dwelling people aged over 65 years falling each year<sup>(1,2)</sup>. Studies have identified many risk factors for falls in older adults<sup>(3,4)</sup>, including intrinsic (e.g., age, gender, cognition, gait and balance, strength, cognition and various diseases) and environmental (e.g., poor lighting, loose rugs and footwear) factors. Of those intrinsic factors, poor gait and balance and muscle weakness are the common modifiable risk factors, with an average of 17% (range 4-39% from 12 studies) of falls in people older than 65 years accounted for by gait/balance disorders or weakness<sup>(5)</sup>. Balance deteriorates with aging and age-related loss of muscle strength is considered an important contributor to decreased balance in older people<sup>(6,7)</sup>.

A significant decline in balance occurs as early as 45 to 55 years old<sup>(8)</sup>, making it important to understand the role of muscle strength in maintaining balance in midlife in order to refine preventative interventions and public health campaigns. However, the association between muscle strength and balance has been examined in only a few cross-sectional studies in young<sup>(9)</sup> and middle-aged<sup>(10-12)</sup> people, with conflicting findings being reported probably due to the small sample size in most studies ( $n < 32$ ). The exception is a study in 1346 middle-aged women reporting that cross-sectionally, greater grip strength was associated with better chair rise performance and the ability to balance on one leg with eyes open for 5 seconds<sup>(12)</sup>.

However, a lack of longitudinal studies in younger people means that the longitudinal relationship between muscle strength and balance remains unclear in this population. It is therefore not known whether developing early interventions aiming to achieve a greater reserve of muscle strength in younger age and to maintain muscle strength over time could be a promising strategy to improve balance in midlife and prevent falls in later life.

Therefore, the aim of this study was to examine whether lower limb muscle strength (LMS) in younger women and changes in LMS over 12 years were independently associated with balance in the same women when middle-aged.

## **Materials and Methods**

### **Participants**

This was an additional 10-year follow-up of a 2-year osteoporosis education randomized controlled trial conducted in 2000 in Southern Tasmania, Australia, with details reported previously<sup>(13)</sup>. Briefly, women aged 25-44 years were randomly selected from the 2000 Tasmanian Electoral Roll. Women were recruited if they were free of the following conditions: previous measurement of bone density, thyroid disease, renal failure, malignancy, or rheumatoid arthritis, a history of hysterectomy or hormone replacement therapies, pregnancy or planning pregnancy within 2 years of study entry, or lactating. At baseline, 470 women were randomly assigned to one of two osteoporosis educational interventions: group education using the Osteoporosis Prevention and Self-management course or an information leaflet. Bone mineral density was measured at the spine and hip at baseline and mean spine and hip T-score used to provide feedback of relative fracture risk as part of the education intervention (higher risk (mean T-score < 0) vs. normal risk (mean T-score ≥ 0)). Ethics approval was obtained from the Tasmania Health and Medical Human Research Ethics Committee and all participants gave written informed consent.

### **Outcomes**

Balance was measured at 12 years by four commonly used clinical tests: the timed up and go test (TUG), the step test (ST), the functional reach test (FRT) and the lateral reach test (LRT). These assess balance performance from either a static or dynamic aspect, and are able to differentiate between fallers and non-fallers in older adults<sup>(8)</sup>. All have been validated in older

women and have a high reliability, with normative values determined in women of the age in our study<sup>(8,14,15)</sup>.

The TUG<sup>(16)</sup> is a test of dynamic steady-state balance and gait. Participants sat in an armchair (45 cm high) with their back against the chair, and then stood without using the arms, walked 3 m, turned, walked back, and sat down. The average time of two trials was used for analysis.

The ST<sup>(17)</sup> measures speed of performing a dynamic stepping task. Participants stood 5 cm from an 8.5-cm-high block positioned against a wall and placed the whole foot of one leg onto the block and returned it to the floor repeatedly as fast as possible for 15 seconds. The number of steps was recorded for both sides and the mean number of steps for each side was used in the analyses.

The FRT measures ability to reach forward with each arm from a bilateral stance position<sup>(14)</sup>. Participants stood with feet a comfortable distance apart behind a line perpendicular and adjacent to a wall. The arm closest to the wall was raised to shoulder height and the position of the knuckle of the middle finger measured<sup>(14)</sup>. Participants leaned forward as far as possible and the position of the knuckle was recorded at the point of furthest reach. The mean of the three trials on each side was used in the analyses.

The LRT measures ability to reach to the side in bilateral stance<sup>(15)</sup>. Participants stood with their backs near but not touching a wall with the heels 10 cm apart. Participants raised both arms to shoulder height while the position of the third finger's tip on the side being measured was marked on the wall. Participants then lowered the arm not being measured and reached sideways as far as possible with the arm being measured. The position of furthest reach was marked and the difference between the two marks calculated. The mean of the three trials on each side was used in the analyses.

## Exposures

Lower limb muscle strength (LMS) was measured to the nearest kilogram using a dynamometer (TTM Muscular Meter, Tokyo, Japan) at baseline and 12 years<sup>(18)</sup>. This test examines isometric strength, predominantly of the quadriceps and hip extensors. The examiner demonstrated the correct technique to the participant before testing. Participants stood on the back of the dynamometer platform, with back straight against a wall and knees flexed to an angle of 115°. They participant held a bar, connected to the dynamometer by a chain, and lifted the bar upwards using maximum force using both legs, with the back and neck straight. The mean of two readings was used in the analyses. 12-year change was calculated by subtracting baseline LMS from 12-year LMS.

## Other measurements

Factors measured at baseline and 12 years included *anthropometric factors*: height measured by a stadiometer (The Leicester height measure, Invicta Plastics Ltd, Oadby, England), weight by a single set of calibrated scales (Heine, Dover NH USA) and body mass index (BMI) calculated (weight/height (kg/m<sup>2</sup>)). A standardised questionnaire was used to collect smoking history (current/former/never), breastfeeding history, number of children, family history of osteoporosis and/or fracture, and previous fractures, current use of oral contraceptive pill (yes/no), education level, employment status, and marital status. *Dietary Calcium intake* was assessed by a short food frequency questionnaire (FFQ), which has been validated against 4 day weighed records and correlates well for estimated calcium intake<sup>(19)</sup>. The calcium content of food categories was determined by Australian food composition tables<sup>(20)</sup>. Physical activity was measured using a validated questionnaire<sup>(28)</sup>, which was modified for Tasmanian conditions and used previously in women of this age. This questionnaire assessed strenuous and light physical activity levels by asking participants how many days in the last 2 weeks they

reported performing at least 20 minutes of strenuous exercise and light exercise, represented by five categories (1 = 0 days, 2 = 1-2 days, 3 = 3-5 days, 4 = 6-8 days, 5 = 9 or more days). At 12 years, prescription medication use was assessed by asking participants to report all medication, prescribed by a doctor that they had taken in the last 2 weeks. Participants were also asked to recall if they had regularly used calcium and vitamin D supplements during the last year, where regular use means taking supplements at least 5 times per week for more than 9 months of the year.

### **Statistical analyses**

Participants' characteristics were presented using mean (standard deviation) or number (%) as appropriate. Univariable and multivariable linear regression were used to separately evaluate associations of baseline and 12-yr change in LMS with balance tests, then baseline and change in LMS were combined in the same model to assess whether these associations were independent of each other. We also calculated the adjusted mean values (standard error) of each balance measure for the lowest and highest quartiles of baseline and change in LMS and the difference in the mean values (95% confidence interval) between the lowest and highest quartiles. We assessed potential confounders using three steps: Step 1: we fitted univariable regression models for each balance outcome and potential confounders (educational level, marital status, employment status, current smoking status and calcium intake at baseline and employment status, menopausal status and calcium intake at 12 years), and retained only those with  $p < 0.20$ . Step 2: we fitted multivariable models including: the exposure of interest; age, weight and height at baseline as compulsory covariates; and the additional confounders identified in step 1. Step 3: additional confounders were retained in the model only if their inclusion changed the estimated coefficient of the exposure of interest by more than 10%.

Scatterplots with regression line were created using adjusted data. Adjusted values for each balance test were calculated by regressing each measured balance test on its confounding factors (except for baseline or 12-yr change in LMS), and then adding the residuals to the mean of each measured balance test. Adjusted baseline and 12-yr change in LMS was calculated by the same approach.

To account for missing follow-up data, we repeated the linear regression analyses using propensity weighting. We used a weighted estimating equation method<sup>(21,22)</sup>, in which data were assumed to be missing at random. Logistic regression models were used to estimate the probability of response using baseline characteristics from the original study (i.e., age, education level, current smoking status and marital status), for which complete data were available.

All analyses were performed in Stata version 12 (Stata Corporation, Texas, USA). A two-tailed p value <0.05 was considered statistically significant.

## **Results**

In the original study, a total of 470 women (64% response rate) aged 25-44 years were recruited at baseline (three women withdrew before bone density was measured). The follow-up period ranged from 8.8 to 12.4 years (mean =11.6 years). At 12 years, 347 (74%) were retained and 342 with full data for LMS at baseline and 12 years were analysed in the present study. **The differences in** baseline characteristics of participants who did and did not complete the follow-up have been previously reported<sup>(23)</sup>. Briefly, women lost to follow-up were younger, had lower levels of education, and were more likely to be current or past smokers, and less likely to be married or in a de facto relationship compared to those who were



retained ( $p < 0.05$  for all). There were no differences in other anthropometric and demographic factors.

Characteristics of study participants are given in Table 1. The mean age of study participants was 38.2 (SD 5.2) at baseline and 49.9 (SD 5.2) years at 12 years. There was, on average, a decline of 17.3 (SD 22.4) kg in LMS from baseline to 12 years.

Table 2 gives the associations of baseline LMS and 12-yr changes in LMS with balance tests, initially unadjusted (Model 1), then adjusted by age, weight, height and education level at baseline (Model 2). Model three gives the independent effects of baseline and change in LMS adjusted for potential confounders (also shown in Figures 1 (A-D) and 2 (A-D)). In unadjusted and unadjusted analyses, baseline LMS was significantly associated with all balance tests while 12-yr change in LMS was only associated with FRT. After adjustment for confounders and with both baseline and change in LMS in the model (Model 3), LMS at baseline was associated with better performance on the TUG [-0.008 sec/kg (95% confidence interval (CI): -0.011, -0.006)], ST (0.044 steps/kg: 0.022, 0.067), FRT (0.057 cm/kg: 0.030, 0.084) and LRT (0.030 cm/kg: 0.012, 0.049) 12 years later. These effect sizes equate to differences from the sample mean of approximately 3.8% for TUG, 6.1% for ST, 3.5% for FRT and 4.1% for LRT for every standard deviation increase in baseline LMS. In model 3, slower decline in LMS was also significantly associated with better performance on TUG (-0.006 sec/kg: -0.009, -0.003), FRT (0.074 cm/kg: 0.044, 0.103) and LRT (0.022 cm/kg: 0.002, 0.043) but there was no association with ST. For change in LMS, these effect sizes equate to differences from the sample mean of approximately 2.5% for TUG, 4.1% for FRT and 2.6% for LRT for every standard deviation decrease in the loss of LMS over 12 years.

In propensity-weighted analyses, the fully adjusted association (Model 3) between 12-yr change in LMS and step test became statistically significant (0.023 steps/kg: 0.009, 0.037) while other associations remained similar (Supplemental Table 1).

Table 3 gives the mean baseline LMS and change in LMS for the lowest and highest quartiles of these variables respectively, together with the adjusted mean balance measures and their absolute and relative differences between highest and lowest quartiles of baseline and change in LMS. There were large differences in LMS between these quartiles (64.2 kg for baseline LMS and 66.9 kg for change in LMS). Women in the highest quartiles of baseline and of change in LMS had better performance in all balance measures, with a percentage difference ranging from 4.5% to 11.3% across different balance measures.

## **Discussion**

This is the first study to examine the longitudinal association of LMS in younger women and change in LMS over time with balance in middle life. Both baseline and 12-yr change in LMS are independently associated with all or most balance tests 12 years later. This suggests that achieving a greater reserve of LMS as younger adults and slowing decline of strength subsequently may both protect women from reduced balance in middle age. These data support the importance of testing whether intervening in early adult life to improve muscle strength are a viable strategy to improve balance and reduce falls in later life. Moreover, women having low LMS in early adulthood or a high rate of loss of LMS may be at higher risk of poorer balance in middle age and potentially benefit most from earlier interventions to improve muscle strength. Thus the potential usefulness of clinical screening of LMS and monitoring its decline in healthy younger women warrants further investigation. Such investigation should include assessment of the feasibility and cost-effectiveness of any such program.

While the effects of LMS and its change per kilogram of muscle strength are modest, they still may be of clinical importance when examined in the context of the very substantial mean change in LMS observed in this population of more than 17 kg over 12 years and differences in baseline and change in LMS between the lowest and highest quartiles of these values (64.2 kg for baseline and 56.9 kg for change). The effect sizes for TUG (3.8% and 2.5%), ST (6.1% for baseline LMS only), FRT (3.5% and 4.1%) and LRT (4.1% and 2.6%) for per standard deviation increase in baseline and change in LMS, respectively are much larger than annualized decline of 0.3% and 1.6% for balance (the standard Romberg test) and 0.19% and 0.25% for gait velocity observed in women aged 50 and 60 years at baseline, respectively<sup>(24)</sup>. The relative differences in balance measures between quartiles of baseline and change in LMS are even more marked, ranging from 4.5 to 11.3% across the different balance measures. Furthermore, the effects of exercise programs on LMS can be substantial. For example, in the elderly, a meta-analysis gave a pooled estimate of the effect of progressive resistance training on leg extensor strength demonstrating a strong effect (SMD 0.68, 95%CI: 0.52 to 0.84)<sup>(25)</sup>, so it is feasible that LMS could be improved to a level that can have a clinically important effect on balance in middle-aged women.

These data support the need for further research to underpin the potential development of approaches to clinical screening of LMS and the monitoring of its decline in healthy younger women. However, before actually recommending such activities in clinical practice, a stronger evidence base that (1) links balance in younger adults to falls in the elderly; (2) determines the effectiveness of interventions to improve balance in younger adults; and (3) demonstrates the feasibility and cost-effectiveness of such screening would be needed.

Although it has long been shown that muscle strength is important for physical functions such as balance, mobility and disability, longitudinal data has been mainly reported in older

populations<sup>(26-28)</sup>, with only one in middle-aged men<sup>(29)</sup>. The finding that both baseline and change in muscle strength are independent predictors of balance in middle age suggests that greater reserve of muscle strength in younger women and the prevention of muscle loss may both be clinically important for achieving better balance in midlife. This is analogous to the confirmation of the importance of the contributions of peak bone mass and bone loss to fracture risk<sup>(30)</sup>, the first reports of which in the 1990's <sup>(31,32)</sup> resulted in a paradigm shift in the approach to osteoporosis prevention towards a greater emphasis on optimising bone acquisition during growth, and maintaining BMD in middle age rather than focussing solely on slowing bone loss in the elderly<sup>(33)</sup>. Modelling of lumbar spine BMD suggests that a 10% improvement in peak bone mass could delay the onset of osteoporosis by 13 years, compared to a 10% increase in the age of menopause (triggering menopausal bone loss) causing a delay of only 2 years<sup>(34)</sup>. Similarly, optimisation of adult LMS could potentially prevent falls and related injuries in old age<sup>(35)</sup>, though direct evidence for the relationship between midlife balance and falls risk in older age is currently lacking. Nonetheless, as for bone, our data suggest that further investigation of the potential improvements in balance from optimising LMS is warranted.

There are effective interventions of improving muscle strength, of which, resistance training has been shown to be effective throughout lifetime<sup>(36-40)</sup>. However, the effectiveness of exercise training for improving balance have to date been rarely examined in younger healthy individuals. We are aware of only one pre-post intervention study showed that both LMS and balance could be significantly improved by strength training in middle-aged women<sup>(11)</sup>. Therefore, more studies of high quality, such RCTs, are needed to confirm the effectiveness of exercise training for improving balance in younger people.

This study has several potential limitations. The balance tests used have not been validated in women of the age studied, mainly because there is no gold standard for balance test in this age group. While we are not aware of data assessing the predictive value in exclusively middle-aged people, balance tests including the TUG and bipedal stance on foam eyes closed, predicted multiple falls in women over 40 years old, 51% of whom were in the age range 40-60 years i.e. not dissimilar to our study population<sup>(41)</sup>. This together with the clear association between all the measures we used and age across the entire age range of 20 – 80 years in a large population based study<sup>(8)</sup>, support the use of these measures in our study as does is the substantial evidence supporting this link in older people (as outlined above). Nonetheless, longitudinal studies specifically in younger adults are still needed to directly confirm the importance of improving balance in younger people to reducing falls risk in later life. However, by its nature, these longitudinal data over decades of follow up are going to be the most challenging and expensive data to obtain. In our view, our data provide important bridging evidence to support undertaking such a study. Although isometric muscle strength is easier to measure and apply in clinic and community settings, isotonic muscle strength has been considered a better predictor of falls in older people<sup>(42)</sup>. However, whether this also applies to younger people is unknown, but studies have shown inconsistent capacity of muscle power and strength in predicting balance in younger adults<sup>(43)</sup>. Therefore, future studies in this age group should consider the difference between isometric and isotonic muscle strength in the capacity of predicting falls risk. As discussed in our previous study, the 64% response rate at recruitment may have led to selection bias<sup>(44)</sup>. However, while the smoking prevalence at baseline in study participants (17%) was lower than the Tasmanian prevalence of daily smoking (29%) in women aged 25 to 44 years in 1998, the distribution of socioeconomic factors like educational levels and the unemployment rate in our study approximate the overall population figures. Another potential limitation is missing data due

to drop-out with women lost to follow-up being younger, less educated, more likely to be former or current smokers, and less likely to be married or in a de facto relationship<sup>(23)</sup>.

However, sensitivity analyses were performed to take into account missing data by using inverse probability weighting and the results were similar. Therefore, the study findings are likely to be generalisable to healthy Caucasian women in this age range. Baseline balance was not measured in our study, so we could not examine the effects of LMS on change in balance. However, it seems likely that balance measurements at 12 years would be a better predictor of falls risk at 12 years than balance 12 years earlier.

In conclusion, this study suggests that both a greater reserve of muscle strength among younger women and slower decline with ageing have independent roles in maintaining healthier balance in middle age. Moreover, baseline and longitudinal changes in LMS may help early assessment of the risk of decreased balance in middle age related to insufficient muscle strength, for whom interventions of improving muscle strength, such as physical exercise and nutritional interventions, may be needed. Randomized controlled trials are required to examine the impact of increased LMS in younger women on balance in midlife.

## **Acknowledgments**

This research was supported by National Health and Medical Research Council (APP1003437) and RACGP/Osteoporosis Australia Bone Health Research Grant. TW was supported by NHMRC/PHCRED Career Development Fellowship (grant number APP102859) and GJ is supported by NHMRC Practitioner Fellowship. MC is supported by an NHMRC Early Career Fellowship. LL is supported by an Arthritis Foundation Australia – Australian Rheumatology Association (AFA–ARA) Heald Fellowship, funded by the Australian Rheumatology Association and Vincent Fairfax Family Foundation; and a NMHRC Early Career Fellowship (Australian Clinical Research Fellowship) (grant number APP1070586). The authors would like to thank all staff and participants involved in this study.

## **Authors' roles**

Study design: FW, TW, GJ and MC. Study conduct: TW and GJ. Data collection and management: TW, FW and GJ. Data analysis: FW and TW. Data interpretation: all authors. Drafting manuscript: FW and TW. Revising manuscript content: all authors. Approving final version of manuscript: all authors. FW takes responsibility for the integrity of the data analysis.

## References

1. Tromp AM, Pluijm SM, Smit JH, Deeg DJ, Bouter LM, Lips P. Fall-risk screening test: a prospective study on predictors for falls in community-dwelling elderly. *J Clin Epidemiol.* 2001;54(8):837-44.
2. Rapp K, Freiberger E, Todd C, et al. Fall incidence in Germany: results of two population-based studies, and comparison of retrospective and prospective falls data collection methods. *BMC Geriatr.* 2014;14:105.
3. Barrett-Connor E, Weiss TW, McHorney CA, Miller PD, Siris ES. Predictors of falls among postmenopausal women: results from the National Osteoporosis Risk Assessment (NORA). *Osteoporos Int.* 2009;20(5):715-22.
4. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. *Maturitas.* 2013;75(1):51-61.
5. Rubenstein LZ. Falls in older people: epidemiology, risk factors and strategies for prevention. *Age Ageing.* 2006;35 Suppl 2:ii37-ii41.
6. Fukagawa NK, Wolfson L, Judge J, Whipple R, King M. Strength Is a Major Factor in Balance, Gait, and the Occurrence of Falls. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences.* 1995;50A(Special Issue):64-7.
7. Puthoff ML, Nielsen DH. Relationships among impairments in lower-extremity strength and power, functional limitations, and disability in older adults. *Phys Ther.* 2007;87(10):1334-47.
8. Isles RC, Choy NL, Steer M, Nitz JC. Normal values of balance tests in women aged 20-80. *J Am Geriatr Soc.* 2004;52(8):1367-72.
9. Muehlbauer T, Gollhofer A, Granacher U. Association of balance, strength, and power measures in young adults. *J Strength Cond Res.* 2013;27(3):582-9.
10. Muehlbauer T, Gollhofer A, Granacher U. Relationship between measures of balance and strength in middle-aged adults. *J Strength Cond Res.* 2012;26(9):2401-7.
11. Holviala JH, Sallinen JM, Kraemer WJ, Alen MJ, Hakkinen KK. Effects of strength training on muscle strength characteristics, functional capabilities, and balance in middle-aged and older women. *J Strength Cond Res.* 2006;20(2):336-44.
12. Hardy R, Cooper R, Aihie Sayer A, et al. Body mass index, muscle strength and physical performance in older adults from eight cohort studies: the HALCyon programme. *PLoS One.* 2013;8(2):e56483.
13. Winzenberg TM, Oldenburg B, Frendin S, De Wit L, Jones G. Effects of bone density feedback and group education on osteoporosis knowledge and osteoporosis self-efficacy in premenopausal women: a randomized controlled trial. *J Clin Densitom.* 2005;8(1):95-103.
14. Duncan PW, Weiner DK, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. *J Gerontol.* 1990;45(6):M192-7.
15. Brauer S, Burns Y, Galley P. Lateral reach: a clinical measure of medio-lateral postural stability. *Physiother Res Int.* 1999;4(2):81-8.
16. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc.* 1991;39(2):142-8.
17. Hill KD, Bernhardt J, McGann AM, Maltese D, Berkovits D. A New Test of Dynamic Standing Balance for Stroke Patients: Reliability, Validity and Comparison with Healthy Elderly. *Physiotherapy Canada.* 1996;48(4):257-62.
18. Dore D, Quinn S, Ding C, et al. Natural history and clinical significance of MRI-detected bone marrow lesions at the knee: a prospective study in community dwelling older adults. *Arthritis Res Ther.* 2010;12(6):R223.
19. Angus RM, Sambrook PN, Pocock NA, Eisman JA. A simple method for assessing calcium intake in Caucasian women. *J Am Diet Assoc.* 1989;89(2):209-14.
20. English R, Lewis J. Nutritional values of Australian foods. Canberra: AGPS. 1991.



21. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med*. 2012;367(14):1355-60.
22. Vansteelandt S, Carpenter J, Kenward MG. Analysis of Incomplete Data Using Inverse Probability Weighting and Doubly Robust Estimators. *Methodology-European Journal of Research Methods for the Behavioral and Social Sciences*. 2010;6(1):37-48.
23. Wu F, Laslett LL, Wills K, Oldenburg B, Jones G, Winzenberg T. Effects of individualized bone density feedback and educational interventions on osteoporosis knowledge and self-efficacy: a 12-yr prospective study. *J Clin Densitom*. 2014;17(4):466-72.
24. Daly RM, Rosengren BE, Alwis G, Ahlborg HG, Sernbo I, Karlsson MK. Gender specific age-related changes in bone density, muscle strength and functional performance in the elderly: a 10 year prospective population-based study. *BMC Geriatr*. 2013;13:71.
25. Latham NK, Bennett DA, Stretton CM, Anderson CS. Systematic review of progressive resistance strength training in older adults. *J Gerontol A Biol Sci Med Sci*. 2004;59(1):48-61.
26. Taekema DG, Gussekloo J, Maier AB, Westendorp RG, de Craen AJ. Handgrip strength as a predictor of functional, psychological and social health. A prospective population-based study among the oldest old. *Age Ageing*. 2010;39(3):331-7.
27. Buchman AS, Wilson RS, Boyle PA, Tang Y, Fleischman DA, Bennett DA. Physical activity and leg strength predict decline in mobility performance in older persons. *J Am Geriatr Soc*. 2007;55(10):1618-23.
28. Visser M, Goodpaster BH, Kritchevsky SB, et al. Muscle mass, muscle strength, and muscle fat infiltration as predictors of incident mobility limitations in well-functioning older persons. *J Gerontol A Biol Sci Med Sci*. 2005;60(3):324-33.
29. Rantanen T, Guralnik JM, Foley D, et al. Midlife hand grip strength as a predictor of old age disability. *JAMA*. 1999;281(6):558-60.
30. Nguyen TV, Center JR, Eisman JA. Femoral neck bone loss predicts fracture risk independent of baseline BMD. *J Bone Miner Res*. 2005;20(7):1195-201.
31. Hansen MA, Overgaard K, Riis BJ, Christiansen C. Role of peak bone mass and bone loss in postmenopausal osteoporosis: 12 year study. *BMJ*. 1991;303(6808):961-4.
32. Riis BJ, Hansen MA, Jensen AM, Overgaard K, Christiansen C. Low bone mass and fast rate of bone loss at menopause: equal risk factors for future fracture: a 15-year follow-up study. *Bone*. 1996;19(1):9-12.
33. Prevention NIHCDPoO, Diagnosis, and T. Osteoporosis prevention, diagnosis, and therapy. *JAMA*. 2001;285(6):785-95.
34. Hernandez CJ, Beaupre GS, Carter DR. A theoretical analysis of the relative influences of peak BMD, age-related bone loss and menopause on the development of osteoporosis. *Osteoporos Int*. 2003;14(10):843-7.
35. Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ. Will my patient fall? *JAMA*. 2007;297(1):77-86.
36. Borst SE. Interventions for sarcopenia and muscle weakness in older people. *Age Ageing*. 2004;33(6):548-55.
37. Lesinski M, Prieske O, Granacher U. Effects and dose-response relationships of resistance training on physical performance in youth athletes: a systematic review and meta-analysis. *Br J Sports Med*. 2016;50(13):781-95.
38. Behringer M, Vom Heede A, Yue Z, Mester J. Effects of resistance training in children and adolescents: a meta-analysis. *Pediatrics*. 2010;126(5):e1199-210.
39. Roelants M, Delecluse C, Goris M, Verschueren S. Effects of 24 weeks of whole body vibration training on body composition and muscle strength in untrained females. *Int J Sports Med*. 2004;25(1):1-5.
40. Newton RU, Hakkinen K, Hakkinen A, McCormick M, Volek J, Kraemer WJ. Mixed-methods resistance training increases power and strength of young and older men. *Med Sci Sports Exerc*. 2002;34(8):1367-75.

41. Nitz JC, Stock L, Khan A. Health-related predictors of falls and fractures in women over 40. *Osteoporos Int.* 2013;24(2):613-21.
42. Skelton DA, Kennedy J, Rutherford OM. Explosive power and asymmetry in leg muscle function in frequent fallers and non-fallers aged over 65. *Age Ageing.* 2002;31(2):119-25.
43. Muehlbauer T, Gollhofer A, Granacher U. Associations Between Measures of Balance and Lower-Extremity Muscle Strength/Power in Healthy Individuals Across the Lifespan: A Systematic Review and Meta-Analysis. *Sports Med.* 2015;45(12):1671-92.
44. Winzenberg T, Oldenburg B, Freuden S, De Wit L, Riley M, Jones G. The effect on behavior and bone mineral density of individualized bone mineral density feedback and educational interventions in premenopausal women: a randomized controlled trial [NCT00273260]. *BMC Public Health.* 2006;6:12.

**Table 1** Characteristics of study participants in younger and middle age (n=342)

Characteristic	
<b>Younger age (Baseline)</b>	
Age (yr.)	38.2 (5.2)
Height (cm)	163.4 (6.3)
Weight (kg)	69.6 (13.4)
Body mass index (kg/m <sup>2</sup> )	26.0 (4.8)
Education, n (%) <sup>a</sup>	
Primary school	1 (0)
High school	102 (30)
High school/college	75 (22)
University, CAE or other tertiary institution	163 (48)
Dietary calcium intake (mg/d)	802 (411)
Current smoking n (%)	43 (13)
Married or de facto, n (%)	260 (76)
Number of children (median) (IQR)	2 (1-2)
Ever smoked, n (%)	156 (46)
Employment status	
0 h/week	43 (13)
≤ 20 h/week	83 (24)
> 20 h/week	216 (63)
Strenuous activity level, median (IQR)	3.0 (1.4)
Family history of fracture, n (%)	120 (35)
History of fracture, n (%)	99 (29)
Lower limb muscle strength (kg)	92.8 (25.4)
<b>Middle age (12 year follow-up)</b>	
Age (years)	49.9 (5.2)
Height (cm)	164.0 (6.1)
Weight (kg)	73.8 (15.8)
Body mass index (kg/m <sup>2</sup> )	27.4 (5.8)
Strenuous activity level, median (IQR)	3.0 (1.4)
Calcium supplement use n (%)	123 (35)
Vitamin D supplement use n (%)	126 (36)
Current use of oral contraceptive n (%) <sup>b</sup>	37 (11)
Current use of estrogens n (%) <sup>b</sup>	24 (7)
History of fracture n (%)	141 (41)
Timed up and go test (seconds)	5.29 (0.68)
Step test (steps)	18.4 (2.6)
Functional reach test (cm)	41.3 (6.2)
Lateral reach test (cm)	18.7 (3.9)
Lower limb muscle strength (kg)	
12 years	75.5 (25.1)
Change from baseline	-17.3 (22.4)

Values are Mean (SD) unless otherwise stated.

<sup>a</sup> primary school (left before the end of grade 10); high school (completed to the end of grade 10); high school/college (completed to the end of grade 12); university, CAE or other tertiary institution.

<sup>b</sup> use in the last two weeks.

**Table 2** Linear regression for associations of baseline and 12-yr change in LMS (kg) with balances at 12 years (n =342)

	Model 1	Model 2	Model 3
	$\beta$ (95% CI)	$\beta$ (95% CI)	$\beta$ (95% CI)
<b>Baseline LMS</b>			
Timed up and go test (seconds)	<b>-0.007 (-0.010, -0.004)</b>	<b>-0.006 (-0.008, -0.003)</b>	<b>-0.008 (-0.011, -0.006)</b>
Step test (steps)	<b>0.025 (0.015, 0.036)</b>	<b>0.036 (0.016, 0.056)</b>	<b>0.044 (0.022, 0.067)</b>
Functional reach test (cm)	<b>0.054 (0.028, 0.080)</b>	<b>0.025 (0.0002, 0.050)</b>	<b>0.057 (0.030, 0.084)</b>
Lateral reach test (cm)	<b>0.031 (0.015, 0.047)</b>	<b>0.022 (0.005, 0.038)</b>	<b>0.030 (0.012, 0.049)</b>
<b>12-yr change in LMS</b>			
Timed up and go test (seconds)	-0.001 (-0.004, 0.002)	-0.001 (-0.004, 0.001)	<b>-0.006 (-0.009, -0.003)</b>
Step test (steps)	0.002 (-0.011, 0.014)	-0.002 (-0.024, 0.017)	0.021 (-0.004, 0.046)
Functional reach test (cm)	<b>0.044 (0.015, 0.073)</b>	<b>0.045 (0.018, 0.071)</b>	<b>0.074 (0.044, 0.103)</b>
Lateral reach test (cm)	0.004 (-0.015, 0.023)	0.006 (-0.012, 0.025)	<b>0.022 (0.002, 0.043)</b>

Bold denotes statistical significance,  $p < 0.05$ .

LMS, lower limb muscle strength.

Model 1, unadjusted.

Model 2, adjusted for age, weight, height and education level at baseline.

Model 3, model 2 + 12-yr change in LMS when baseline LMS was the exposure of interest and vice versa.

**Table 3** Difference in balance measures between the lowest and highest quartile of baseline and change in LMS

	Lowest quartile	Highest quartile	Absolute difference (95% CI)	Percentage difference (%) <sup>b</sup>
Baseline LMS (kg) <sup>a</sup>	62.1 (11.0)	126.3 (14.8)		
Timed up and go test (seconds)	5.56 (0.06)	5.03 (0.06)	<b>-0.531 (-0.719 to -0.342)</b>	9.5
Step test (step)	17.39 (0.28)	19.36 (0.28)	<b>1.97 (1.15 to 2.78)</b>	11.3
Functional reach test (cm)	39.80 (0.63)	43.19 (0.63)	<b>3.39 (1.54 to 5.24)</b>	8.5
Lateral reach test (cm)	17.76 (0.44)	19.35 (0.44)	<b>1.58 (0.30 to 2.87)</b>	8.9
12-year change in LMS (kg) <sup>a</sup>	-45.7 (12.8)	11.2 (11.2)		
Timed up and go test (seconds)	5.39 (0.06)	5.15 (0.06)	<b>-0.233 (-0.425 to -0.054)</b>	4.5
Step test (step)	17.88 (0.27)	18.74 (0.27)	<b>0.86 (0.07 to 1.65)</b>	4.8
Functional reach test (cm)	39.02 (0.63)	43.30 (0.62)	<b>4.28 (2.47 to 6.09)</b>	11.0
Lateral reach test (cm)	17.88 (0.43)	19.12 (0.43)	1.24 (-0.01 to 2.49) <sup>c</sup>	6.9

LMS, lower limb muscle strength; CI, confidence interval.

Values are adjusted mean (standard error) unless otherwise stated, adjusted for age, weight, height, education level and baseline or change in LMS as appropriate.

Bold denotes statistical significance.

<sup>a</sup> values are unadjusted mean (standard deviation).

<sup>b</sup>  $100 \times$  absolute difference/mean in lowest quartile.

<sup>c</sup>  $p = 0.052$ .

## Figure legends

Figure 1, A-D, Scatter plots and linear regression lines for associations of baseline LMS and balance tests after 12 years. Linear regression lines are from models adjusted for the same confounders as in Table 2. Higher values of timed up and go test represent poorer performance whereas higher values of all other tests represent better performance.

Figure 2, A-D, Scatter plots and linear regression lines for associations of 12-yr change in LMS and balance tests after 12 years. Linear regression lines are from models adjusted for the same confounders as in Table 2. Higher values of timed up and go test represent poorer performance whereas higher values of all other tests represent better performance.