A systematic review of one-legged balance performance and falls risk in community-dwelling adults

Joanna M Blodgett [a*]; Jodi P Ventre[bc]; Richard Mills[b]; Rebecca Hardy[d]; Rachel Cooper[b]

[a] Institute of Sport, Exercise & Health, Division of Surgery & Interventional Science, University College London, 170 Tottenham Court Road, W1T 7HA, London, UK
[b] Department of Sport and Exercise Sciences, Musculoskeletal Science and Sports Medicine Research Centre, Manchester Metropolitan University, Oxford Road, M15 6BH, Manchester, UK
[c] Department of Psychology, Health, Psychology and Communities Research Centre, Manchester Metropolitan University, Bonsall Street, M15 6GX, Manchester, UK
[d] CLOSER, Social Research Institute, University College London, 55-59 Gordon Square, WC1H 0NU, London, UK

* Corresponding Author:
Joanna M Blodgett
Institute of Sport, Exercise & Health
Division of Surgery & Interventional Science
University College London
170 Tottenham Court Road
W1T 7HA
London, UK
Email: joanna.blodgett@ucl.ac.uk

Abstract

Objective: The aim of this systematic review was to synthesise all published evidence on associations between one-legged balance performance and falls. **Methods**: Medline, EMBASE, CINAHL and Web of Science were systematically searched (to January 2021) to identify peer-reviewed, English language journal articles examining the association between one-legged balance performance and falls in community-dwelling adults. **Results**: Of 4 310 records screened, 55 papers were included (n=36 954 participants). There was considerable heterogeneity between studies including differences in study characteristics, ascertainment of balance and falls, and analytical approaches. A meta-analysis of the time that individuals could maintain the one-legged balance position indicated that fallers had worse balance times than non-fallers (standardised mean difference: -0.29(95%CI:-0.38,-0.20) in cross-sectional analyses; -0.19(-0.28,-0.09) in longitudinal analyses), although there was no difference in the pooled median difference. Due to between-study heterogeneity, regression estimates between balance and fall outcomes could not be synthesised. Where assessed, prognostic accuracy indicators suggested that one-legged balance was a poor discriminator of fall risk; for example, 5 of 7 studies demonstrated poor prognostic accuracy (Area Under the Curve <0.6), with most studies demonstrating poor sensitivity. **Conclusions**: This systematic review identified 55 papers that examined associations between balance and fall risk, the majority in older aged adults. However, the evidence was commonly of low quality and results were inconsistent. This contradicts previous perceptions of one-legged balance as a useful fall risk tool and highlights crucial gaps that must be addressed in order to translate such assessments to clinical settings.

Keywords: one-legged balance; falls; systematic review; community-dwelling

1.0 INTRODUCTION

Falls are a leading cause of injury, functional impairment, and death in older adults(Ungar et al., 2013). Globally, an estimated 28 to 50% of individuals over the age of 65 reported a fall in the past year(Soriano et al., 2007; WHO, 2007). Falls have substantial impacts at both individual and population levels. A recent Global Burden of Disease study estimated that falls resulted in 16.7 million years of life lost, 19.3 million years lived with disability and 35.9 disability-adjusted life years(James et al., 2020). This is consistent with World Health Organisation reports suggesting that falls are the leading cause of injury-related death in adults aged ≥ 65 years(WHO, 2007) and estimating that annually, falls cause 684,000 deaths with over 37 million falls severe enough to warrant medical attention(WHO, 2021). Annual medical costs associated with falls are estimated to be $50 billion in the USA and £2.3 billion in the UK and continue to rise(Florence et al., 2018; NICE guideline, 2013). There is emerging evidence that midlife may represent an important period for fall-related interventions, with pooled analysis demonstrating that fall prevalence is already significant in adults aged 40 to 64 years, ranging from 8.7% to 31.1% (Peeters et al., 2019; Peeters et al., 2018).

Successful fall prevention strategies must consider effective screening tools, targeted interventions that mitigate risk factors, and modification of home or community environments to reduce extrinsic hazards (Dellinger, 2017; Hopewell et al., 2018). Of the many risk factors studied, history of falls and balance or gait impairments have been identified as the two strongest predictors of future falls(Ganz et al., 2007). Given the role of balance in maintaining postural stability, improving balance ability in older adults is frequently a target of falls prevention interventions (Sherrington et al., 2019). Further, balance assessments are commonly used in research and clinical settings as a prognostic tool to identify those at higher risk of falling(Springer et al., 2007; Vereeck et al., 2008). Balance tests in these settings are highly heterogeneous. For example, some balance tests use performance-based measures such as the one-legged stand or functional reach test, while others rely on cumulative, subjective measures such as the Tinetti Assessment Test or the Berg Balance Scale, which consist of 9 and 14 balance-related tasks, respectively, each scored on 3 to 5 point scales(Mancini and Horak, 2010).

Previous systematic reviews have examined the utility of single (Barry et al., 2014; Lima et al., 2018; Moore and Barker, 2017; Okubo et al., 2021; Rosa et al., 2019) or multiple (Gates et al., 2008; Kozinc et al., 2020; Lusardi et al., 2017; Okubo et al., 2021; Power et al., 2014) balance measures in predicting falls. These reviews commonly focused on older adults (≥60 years) and had broad inclusion criteria; for example, studies from any setting (e.g. clinical vs community-dwelling) or that used any balance measure were often eligible for inclusion. No review has focused exclusively on the one-legged balance test and reviews of multiple balance measures reported conflicting evidence on one-legged balance test and fall risk (Kozinc et al., 2020; Lusardi et al., 2017; Power et al., 2014). In addition, the broad search terms used to capture multiple measures of balance in a single review did not identify all studies examining one-legged balance and falls.

The one-legged balance test is one of the most commonly used balance tests and is widely considered to be cost-effective and feasible in both clinical and research settings(Bohannon, 2006; Jonsson et al., 2004; Mancini and Horak, 2010; Michikawa et al., 2009; Springer et al., 2007). Proponents of the test suggest that it should be implemented into primary care to help identify individuals at higher risk of falling and other poor health outcomes(Kozinc et al., 2020; Michikawa et al., 2009; Nickelston, 2014), emphasising a clear need to systematically review and synthesise the evidence on one-legged balance performance and fall risk. To address this gap and provide a robust summary of available evidence, we undertook a systematic review and meta-analyses to synthesise all published evidence of associations between one-legged balance performance and fall risk in community-dwelling adults. We hypothesised that there would be consistent evidence of an association between better one-legged balance performance and lower fall risk.

2.0 METHODS

This systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines(Moher et al., 2009) and the study protocol was registered with PROSPERO (CRD42020160413)(Blodgett et al., 2020c).

2.1 Eligibility criteria

Studies published in peer-reviewed journals were eligible for inclusion if they examined the association between one-legged balance performance and any fall outcome in a community-dwelling sample. Studies were excluded if they were: published in a non-English language journal; systematic reviews or intervention studies; or if they considered a specific clinical sample (e.g. those with Parkinson's disease).

2.2 Search strategy

We searched Medline, EMBASE, CINAHL and Web of Science for all available articles from inception to January 2021. Two distinct search arms were combined using the Boolean term "AND". All possible synonyms and truncations of one-legged balance synonyms (e.g. single leg, flamingo or unipedal stand) comprised one arm, while "fall" with any truncation constituted the other. See Appendix A for the complete search strategy. Reference lists of all included articles were independently searched by two authors to identify additional studies.

2.3 Study selection

All articles were uploaded into Mendeley and Rayyan(Ouzzani et al., 2016), which were used to remove duplicates and manage the two-stage screening process. In both the title-abstract and full-text screening stage, two authors (JB, JV or RM) independently screened all potential papers for inclusion. In the full-text screening stage, each author recorded the reason for exclusion following a hierarchical list of five criteria (outlined in Figure 1). Discrepancies in screening decision or exclusion rationale were resolved through discussion between authors.

2.4 Data extraction

Two authors (JB, JV, RM or RC) independently extracted data and any conflicts were resolved through discussion. For all included papers, the following data were extracted using a standard proforma in Google Forms (see Appendix B): demographic characteristics (country, study design, exclusion criteria, sample size, sex and age), one-legged balance (assessment protocol details), falls (definition, prevalence, data collection protocol, outcome type), statistical methods, and effect estimates. WebPlotDigitizer was used to extract data that were presented in graphs and not tables(Rohatgi, 2020). A modified version of the Newcastle-Ottawa Risk of Bias Scale was used to appraise the quality of each included study (see Appendix B, part 6). Scores ranged from zero (lowest quality) to seven (highest quality). Any discrepancies in scores were discussed and resolved by authors.

2.5 Narrative synthesis

Narrative synthesis of study characteristics, one-legged balance measurement and falls measurement was first conducted following established guidelines(Popay et al., 2006). Results are presented by fall outcomes: any fall (0 vs 1+ fall), recurrent falls (0-1 vs 2+ falls) or injurious falls (non-injurious or 0 injurious falls vs 1+ injurious falls). For associations between one-legged balance and falls, meta-analyses were conducted where there were comparable estimates from three or more studies and a narrative synthesis of estimates was conducted if meta-analyses were not possible. It was decided *a priori* that estimates could not be synthesised in meta-analyses if there were differences in temporality (e.g. cross-sectional or longitudinal), model adjustment (e.g. unadjusted or adjusted) or balance dichotomisation (e.g. ≤5s, ≤30s, etc.).

Where studies presented multiple estimates for an association (e.g. balance times for both legs, best and average balance trials, multiple balance cut-points or results for balance with eyes open and closed), a single result was used in the main analysis although all results are presented in Appendix C. The result provided in the main analysis is selected based on comparability with other papers (e.g. common characteristics as demonstrated in the initial narrative synthesis) and completeness of data (e.g. estimates, error terms). Where studies presented associations for multiple fall outcomes, effect estimates for each outcome were considered in each relevant section. To maximise comparison of results between studies, odds ratios (OR) and prognostic indicators (i.e. sensitivity, specificity, positive predictive value, negative predictive value) were calculated from proportions and sample sizes, where possible.

2.6 Meta-analyses

Meta-analysis of median differences was conducted using the *metamedian* package in R, which provides an estimate for the weighted pooled difference of median balance times between fallers and non-fallers(McGrath et al., 2020). Meta-analysis of standardised mean difference (SMD) in balance time between fallers and non-fallers was conducted using the package *meta* in R to calculate Hedge’s g(Balduzzi et al., 2019; Lakens, 2013). Hedge’s g (i.e. SMD) is a measure of the effect size and is calculated as the difference in mean balance times between groups divided by the standard deviation of the combined sample. Due to difference in the length of balance trials between studies, raw mean difference times were not appropriate due to dissimilar scales and ceiling effects. Where standard errors (SE) were missing or could not be calculated from available information, inclusion of studies in the meta-analysis was maximised using a prognostic imputation method to impute SE(Ma et al., 2008). Random-effects models were used to estimate and compare SMDs by cross-sectional and longitudinal subgroups. As a supplementary analysis, we further stratified by age group (<75 years, ≥ 75 years). The I^2 statistic was considered as the indicator of between-study heterogeneity, where 25%, 50% and 75% suggest low, moderate and high heterogeneity, respectively(Higgins et al., 2003). Finally, publication bias was examined using the Egger test and visual inspection of a funnel plot(Sterne and Egger, 2005). To ensure no single study was driving the result, a sensitivity analysis repeated the Egger test multiple times, removing each study in turn. Due to heterogeneity outlined in section 2.5, meta-regression was not possible for any fall outcome. All meta-analyses were conducted in R Studio version 1.2.5.

3.0 RESULTS

Our database searches identified a total of 4,310 unique records. After the two-stage screening and additional papers identified via the reference list search, a total of 55 papers are included in the review(Andresen et al., 2006; Ansai et al., 2016; Arai et al., 2020; Beauchet et al., 2010; Bergland and Wyller, 2004; Blain et al., 2021; Bongue et al., 2011; Briggs et al., 1989; Buatois et al., 2006; Buatois et al., 2010; Cho and Kamen, 1998; Choy et al., 2008; Choy et al., 2007; Crenshaw et al., 2020; de Rekeneire et al., 2003; Delbaere et al., 2010; Depasquale and Toscano, 2009; Ek et al., 2019a; Ek et al., 2019b; El-Sobkey, 2011; Eto and Miyauchi, 2018; Gerdhem et al., 2005; Hasegawa et al., 2019; Hashidate et al., 2011; Heitmann et al., 1989; Ikegami et al., 2019; Jalali et al., 2015; Kwan et al., 2011; Lim et al., 2016; Lin et al., 2004b; MacRae et al., 1992; Mahoney et al., 2019; Moreira et al., 2017; Muir et al., 2010; Mulasso et al., 2017; Nevitt et al., 1989; Niam and Wee, 1999; Park et al., 2020; Porto et al., 2020; Rossat et al., 2010; Sampaio et al., 2013; Shimada et al., 2009; Shimada et al., 2011; Shin et al., 2012; Shinohara et al., 2020; Swanenburg et al., 2013; Thomas and Lane, 2005; Tinetti et al., 1988; Toulotte et al., 2006; Vellas et al., 1998; Vellas et al., 1997; Welmer et al., 2017; Yamada and Ichihashi, 2010; Yamada et al., 2012; Yamada et al., 2020)(see Figure 1). The 55 papers use data from 51 study samples, with multiple papers using data from the Swedish National Study on Ageing and Care in Kungsholmen(Ek et al., 2019a; Ek et al., 2019b; Welmer et al., 2017), the Albuquerque Falls Study(Vellas et al., 1998; Vellas et al., 1997) and an unnamed French cohort(Beauchet et al., 2010; Bongue et al., 2011). Characteristics of all studies are presented in Table 1 and Appendix C.

3.1 Description of studies, balance and falls

3.1.1 Study characteristics

Thirty papers assessed cross-sectional associations between balance and falls, 22 assessed longitudinal associations (follow-up range: 12 months to 10 years) and 3 assessed both. Studies were conducted in sixteen different countries (see Table 1), with the most common continents being Asia (n=19), North America (n=12) and Europe (n=12). Sixteen studies used data from previously established cohorts and four were case-control studies. A total of 36,954 individuals were included across the 51 study samples, with individual sample sizes ranging from 16 to 7,463. Eight studies considered women-only samples, while the remaining 43 considered both men and women. In mixed-sex studies, the overall proportion of women was 58.6% and ranged from 30% to 84.4%. The mean age, where reported, ranged from 55 to 81.5 years and the most commonly studied age group was aged ≥65 years (n=26). The mean and median study quality scores on the Newcastle-Ottawa Risk of Bias Scale were both 4, with a range from 1 (lowest quality) to 7 (highest quality); scores for each individual item are provided in Appendix D.

3.1.2 Ascertainment of one-legged balance

As some of the papers reporting on the same study population provided different descriptions of one-legged balance, methods for all 55 papers are summarised below. Most papers recorded continuous balance time (n=44), 10 studies collected a binary measure (e.g. <5 vs ≥5s) and the final paper recorded the number of times the participant's foot touched the ground during a continuous 30 second trial. The most common lengths of the continuous trials were 30 (n=14) and 60 (n=9) seconds, with a range of 10 to 120 seconds; ten studies did not report the maximal time. Continuous balance times were analysed in 31 papers, 22 used distinct categorical or binary cut-points and 1 paper analysed both continuous and binary balance times. Fifteen different cut-points were used to create distinct binary groups; the most common was <5 or ≥5 seconds (n=8).

The number of balance trials ranged from a single trial to 24 trials. Of the 36 papers that conducted multiple trials, different strategies were used to select the balance time for analysis; this included the best time (n=18), worst time (n=1) or average time (n=8). The others did not specify which was used or analysed multiple balance times. Eight papers conducted both eyes open and eyes closed trials, 25 conducted eyes open only and 22 did not describe whether eyes were open or closed. Similarly, 20 papers conducted trials on each leg, 23 studies instructed individuals to use their dominant or preferred leg only, one study used the non-dominant leg and the remaining 11 studies did not provide a description. Finally, the majority of papers did not provide details of instructions on the body position required in protocols (see Appendix C).

3.1.3 Ascertainment of falls

Thirty-six studies assessed falls retrospectively (e.g. fall in last 12 months), thirteen prospectively and two studies measured falls both retrospectively and at follow-up. Of the 38 retrospective fall assessments, 22 used self-reported questionnaires, 15 collected data in interviews and one was based on clinician referral. Prospective collection of falls data included diary or post card submission (n=4), regular phone calls (n=5), linked health records (n=1), postal questionnaires (n=1) and five studies combined diary or postcard submissions with phone calls.

As papers that used the same sample examined different follow-up periods and fall outcomes, summary characteristics are, once again, provided at the paper (n=55) rather than study level. Twelve months was the most frequent time period for fall reporting across both prospective and retrospectively collected data (n=41), followed by 2 years (n=5); the remaining 9 studies each had a distinct follow-up period (range: 3 months to 10 years). Eight studies examined multiple fall outcomes. The most common outcome was any fall (e.g. 0 vs 1+ fall; n=38) followed by recurrent falls (e.g. 0-1 vs 2+ falls; n=7) and injurious falls (e.g. no falls/non-injurious falls vs any injurious falls; n=7). Additionally, eight studies considered the number of falls, either continuously (n=4) or in categories (e.g. 0,1,2+ falls; n=4) and one study considered an aggregate outcome of 2+ non-injurious falls or 1+ injurious fall.

The prevalence of falls ranged from 11.0% to 71.2% (median: 28.9%). Many papers described their definition of a fall (n=36), but 19 did not. Of the 36 papers that provided a falls definition, ten created or adapted their own. Exact phrasing was taken from six existing definitions and was most frequently attributed to the Kellogg Working Group(1987) (n=8) or Tinetti(Tinetti et al., 1988) (n=7) (see Appendix E for falls definitions).

3.2 Any fall (no fall vs 1+ falls)

3.2.1 Median differences

Given the skewed distribution of one-legged balance times, the assumption of normally distributed data needed for parametric tests (e.g. t-tests) is not met, indicating that non-parametric tests (e.g. Mann Whitney U tests) are more appropriate(Nahm, 2016). None of the 8 studies (range: n=30 to 213) that used the Mann Whitney U test found a statistically significant difference in balance times in fallers and non-fallers. A meta-analysis of the four studies that provided median balance times, using the median of the difference of medians method(McGrath et al., 2020), provided further support for no difference between groups (1s (95% CI: -1.2,8.9); see Table 2).

3.2.2 Mean differences

Most studies ignored the non-normal distribution of one-legged balance times, with 15 cross-sectional studies and 9 longitudinal studies presenting mean (SD) balance times in fallers and non-fallers (total n=6 894 across all studies). Meta-analyses suggested that fallers had lower mean balance times than non-fallers (SMD= -0.29 (95% CI: -0.38,-0.20)) in cross-sectional analyses and a similar pattern was seen in longitudinal analyses (SMD= -0.19 (-0.28,-0.09)) (see Figure 2). The SMD was smaller for longitudinal associations, although the test for subgroup differences did not reach statistical significance (p=0.09). Estimated heterogeneity in study outcomes in these meta-analyses was low for both cross-sectional (I^2=14% (0,50%)) and longitudinal (I^2=0% (0,60%)) analyses. Visual inspection of the funnel plot and Egger test (p=0.05) suggested that there may be minimal publication bias (see Appendix F); however, this was primarily driven by Cho and Kamen(1998; n= 16) and Hashidate et al.(2011; n=30)(p=0.16 when removed), although there was no impact on the cross-sectional SMD when Cho and Kamen(1998) and Hashidate et al. (2011) were removed from the meta-analysis (-0.28 (95% CI: -0.37,-0.19)).

When stratified by age, there was evidence to suggest that associations were stronger in younger individuals (see Appendix G). In longitudinal analyses, the SMD in favour of non-fallers was larger in studies with a mean age <75 years (SMD= -0.30 (-0.46,-0.15)) compared to those with a mean age ≥ 75 years ((SMD= -0.13 (-0.25,-0.01)). In cross-sectional studies, there was strong overlap in 95% confidence intervals of the SMD of both age groups; it is noteworthy that 7 of 10 studies with a mean age <75 years found a significant association ((SMD= -0.26 (-0.37,-0.16)) compared to just 1 of 4 studies with a mean age≥ 75 years ((SMD= -0.42 (-0.71,-0.13)).

3.2.3 Regression estimates

Meta-analyses of regression outcomes for any fall outcome were not possible due to heterogeneity in temporality, model adjustment and balance dichotomisation, as outlined in section 2.5. Estimates for risk of any fall are presented in Figure 3, with a detailed table of all estimates and study details in Appendix H. Patterns of association were similar across estimate type (e.g. OR per 1s, OR per low balance cut-point, relative risk (RR) per low balance cut-point) and across cross-sectional and longitudinal models. In unadjusted models, poorer balance performance was associated with increased risk of a fall in seven of ten studies, with three studies reporting non-significant results (Figure 3A-C). Two additional studies, Vellas et al.(1998) and Blain et al.(2021), reported positive associations in men but no associations in women (Figure 3A and 3B). Significant odds ratios in unadjusted models ranged from 1.5 (1.2,1.8) for those with a balance time <12.7s(Bongue et al., 2011) to 8.40 (1.10,64.26) in those with a balance time <55.4s(Eto and Miyauchi, 2018)(Figure 3B).

In adjusted models, most studies (n=9/12) reported no association between balance time and falls. The most commonly included covariates were age, sex, body size and comorbidities; covariates for each model are detailed in Appendix H. Weak associations remained in three studies; a one second increase in balance time was associated with lower risk of falling in two cross-sectional studies(Hasegawa et al., 1989; Moreira et al., 2017), while Muir et al.(2010) reported that those who balanced for <10s had a 1.58 (1.03,2.41) times higher risk of falling after a 12-month follow-up. Only five studies provided unadjusted and adjusted estimates, with the adjusted association remaining in Muir et al.(2010) only (Figure 3C).

3.3 Recurrent falls (0-1 fall vs 2+ falls)

Two studies compared median or mean balance times of recurrent fallers; Porto et al. (2020) reported no difference in mean balance time between single fallers and recurrent fallers (19.1s ±10.4 vs 18.2s ±10.2; p=0.84), while Thomas and Lane (2005) reported lower median balance times in recurrent fallers (0.43s (interquartile range: 1.57) compared to single fallers and non-fallers (2.71s (2.59); p<0.05). All other studies that examined recurrent falls used regression models, with sample sizes ranging from 30(Thomas and Lane, 2005) to 7 643(Rossat et al., 2010). In unadjusted models, six of nine studies reported an association between lower balance time and higher risk of falling two or more times (Figure 4, Appendix H), with no association in the remaining studies(Beauchet et al., 2010; Buatois et al., 2006; Swanenburg et al., 2013). OR estimates ranged from 1.6 (1.2, 2.2) in those who maintained balance for <2s to 15.22 (1.72,133.95) in those who balanced for <1.02s(Thomas and Lane, 2005).

Similar to above, comparison of unadjusted and adjusted estimates was possible in three studies (Figure 4). In the first of these by Jalali et al.(2015), those with low balance time (≤12.7s) were more likely to fall multiple times than those with better balance (>12.7s) (unadjusted OR: 8.54 (95% CI: 4.86,14.99)); this association attenuated to 3.71 (95% CI not reported) after adjustment for age, body mass index, diabetes, functional reach and the Romberg test. In another study, Nevitt et al.(1989) found that the association between balance (<2s) and falls in an unadjusted model was fully attenuated after adjustment for race, fall history, comorbidities and other physical performance tests. Finally, Rossat et al.(2010) reported that low balance time (≤5s) was associated with increased risk of recurrent falls even after adjustments (unadjusted Incident Rate Ratio (IRR) 1.85 (1.67,2.05); adjusted IRR 1.55 (1.39,1.73), adjusted for age, sex, medications, cognitive scores and the sit to stand test).

3.4 Injurious falls (non-injurious or 0 injurious falls vs 1+ injurious falls)

There was inconsistent evidence of an association between balance times and injurious falls in the eight papers that assessed this (Appendix H). Two studies reported no associations between balance and injurious falls in both unadjusted and adjusted models(Andresen et al., 2006; Muir et al., 2010), while another reported that those with balance times in the bottom 50% of the sample had 2.4 (1.1, 5.2) times the odds of an injurious fall(Bergland and Wyller, 2004). Vellas et al. 1998 reported higher risk of injurious falls in women (RR: 2.97 (1.86,4.74)) with low balance time but not men (1.79 (0.78,4.15)).

Finally, using Swedish cohort data, Ek et al.(2019a;2019b) and Welmer et al.(2017) reported associations between low balance time (<5s) and increased risk of injurious falls as measured by linked hospital data in 17 different models (Appendix H). Here, associations were similar in men and women(Ek et al., 2019a; Ek et al., 2019b), but weakened with longer periods of follow-up (e.g. from 3 years to 10 years)(Ek et al., 2019b; Welmer et al., 2017). Estimates remained after adjustment for age and education (Ek et al., 2019b; Welmer et al., 2017), while adjustment for previous history of falls, activities of daily living and grip strength often attenuated the estimates(Ek et al., 2019b; Welmer et al., 2017). Appendix H outlines the 17 models, which considered sex-stratification, multiple follow-up periods and inclusion of different covariates.

3.5 Other results of relevance

3.5.1 Prognostic accuracy of the one-legged balance test

Seven studies provided estimates on the prognostic accuracy of the one-legged balance test. Sensitivity, specificity, positive predictive values and negative predictive values could be calculated from sample size and proportions in an additional 11 studies. Similar to the regression estimates, we were unable to conduct a meta-analysis due to variability in cut-points and fall outcome type. There was substantial variability in prognostic accuracy estimates (see Table 3), however most papers reported higher specificity (range: 46.2-90.3%) than sensitivity (16.7-83.5%) and higher negative predictive values (range: 63.4-95.1%) than positive predictive values (range: 12.1-82.4%). Notably, negative predictive values were also higher when considering recurrent falls compared to any fall. When several cut-points were used within the same study sample(Beauchet et al., 2010; Bongue et al., 2011), higher cut-points (e.g. <7.6s or <12.7s) had greater sensitivity but lower specificity compared to a lower cut-point (e.g. <5s). Finally, the area under the curve (AUC) varied from 0.527(Lin et al., 2004b) to 0.766(Depasquale and Toscano, 2009), but was in the range considered as failed discrimination (i.e. 0.5 to 0.6)(Li and He, 2018) for five of seven studies.

3.5.2 Differences in results by balance test conditions

Five studies reported no differences when comparing associations of balance times under eyes open and eyes closed conditions with falls (Briggs et al., 1989; Choy et al., 2007; Heitmann et al., 1989; Shin et al., 2012; Toulotte et al., 2006), however two studies with small sample sizes reported contradictory associations. Cho and Kamen (1998) reported that mean balance times with eyes open were higher in non-fallers compared with fallers, but that no difference was found for the eyes closed condition. Conversely, El-Sobkey et al.(2011) reported a higher odds of falling in those with low balance time with eyes closed and no association with eyes open times. Other variations in balance protocol did not impact greatly on findings; for example, similar associations were found when the following were considered: right or left stance leg(Ansai et al., 2016; Choy et al., 2008; Moreira et al., 2017), better or worse stance leg(Kwan et al., 2011; Shinohara et al., 2020) or when the first or best trial was used (Heitmann et al., 1989).

3.5.3 Results not captured above

Associations between balance and falls identified in five studies could not be included in the syntheses above as they operationalised balance or falls in a non-standard way that limited comparability or they did not provide sufficient study details to interpret the estimates. Further details on these studies are provided in Appendix I. Briefly, a study by Toulotte et al.(2006) provided support for better balance in non-fallers compared with fallers, a study by de Rekeniere et al.(2003) reported no difference in balance between fallers and non-fallers, and the remaining three studies reported inconsistent or uninterpretable findings(Choy et al., 2008; Choy et al., 2007; Ikegami et al., 2019).

4.0 DISCUSSION

4.1 Main findings

In a systematic review of published studies, we identified 55 papers that had examined the association between one-legged balance performance and fall risk in community-dwelling adults, with the majority of samples aged 65+. Although there was inconsistency in findings, there was some evidence to suggest that non-fallers had better balance times than fallers and that lower one-legged balance time was more strongly associated with increased risk of recurrent falls than any fall. However, studies were often of low quality, had a cross-sectional design and considered unadjusted models only. Many studies assessed balance performance after the fall recall period (e.g. cross-sectional design) and thus reverse causality is likely to explain some of this association. Where adjusted models were presented, results suggested that associations were largely explained by confounders. Additionally, prognostic accuracy of the one-legged test was very poor. Thus, the findings of the review crucially highlight the lack of high-quality empirical evidence to support the use of the one-legged balance test as both a screening tool in clinical settings and as an assessment of fall risk in research settings. This finding has very important implications as it cautions against the premature translation of the one-legged balance test into clinical settings. With low quality and inconsistent evidence, there is an urgent need for better, methodologically robust epidemiological evidence in this area.

4.2 Critical appraisal of studies

Due to considerable between-study heterogeneity in sample characteristics, temporality of associations, and measurement and operationalisation of balance and falls, results should be interpreted with caution. A key challenge in interpreting the findings of included studies is that 33 of the 55 papers examined cross-sectional associations between balance performance and falls within the previous 3 to 25 months. As the balance assessment occurred after the fall reporting period, associations identified may, at least partially, be explained by reverse causality. This is plausible as falls have been shown to precipitate mobility impairment, contribute to fear of falling and lead to declining activity levels(Boyd and Stevens, 2009; Stalenhoef et al., 2002); each of which has detrimental effects on balance ability.

Another key challenge for synthesis and interpretation of estimates was the fact that the distribution of one-legged balance times was overlooked in many study analyses. Skewed distribution of balance times is common as one-legged balance performance tests are vulnerable to both floor and ceiling effects depending on the sample age and complexity of the protocol(Bergquist et al., 2019; Blodgett et al., 2020b; Choi et al., 2014; Morioka et al., 2012). The most commonly reported estimate (n=24 studies) was the difference in mean balance times between fallers and non-fallers, however this comparison does not meet the key assumption of normality required for a parametric test (i.e. t-test)(Bridge and Sawilowsky, 1999; Vickers, 2005). Although we present meta-analyses of these results, the results of the SMD approach could be driven by exceedingly low or high performing individuals, rather than the sample as a whole, and must be interpreted with caution.

Less than a third of papers considered confounding. Where studies did adjust for covariates, these adjustments explained most of the associations between balance and falls. No study provided individual stages of adjustment nor considered if covariates acted as confounders or mediators of the balance-fall associations, which is essential to understand the underlying mechanisms of association between balance and falls.

4.3 Potential sources of heterogeneity

There was high heterogeneity between studies in terms of sample characteristics, ascertainment of balance and falls and reporting of results across studies. As such, we were largely unable to examine how balance-falls associations may differ across sample characteristics (e.g. country, sex). For example, only five papers considered sex differences in their analyses(Blain et al., 2021; Ek et al., 2019a; Ek et al., 2019b; Lim et al., 2016; Vellas et al., 1998), despite conceivable sex differences indicated by better one-legged balance in men and greater prevalence of falls in women(Blodgett et al., 2020a; Cooper et al., 2011; Overstall et al., 1977; Peeters et al., 2018; Springer et al., 2007). Similarly, all but one study (Lim et al., 2016) examined associations across the full age range of their sample. Although stratification of the standardised mean difference meta-analysis by mean age of sample (<75, ≥ 75 years) suggested that associations were stronger in younger adults, there remained substantial heterogeneity in the age range of each sample. Further investigation of age differences within the same sample using homogenous protocols is required.

Differences in balance testing protocols may also partially explain inconsistent findings. The majority of studies did not state the starting position or the criteria that ended the balance trial, despite important factors such as movement in the arms, legs and eyes that contribute to balance performance(Boström et al., 2018; Scholz et al., 2012). As upper body movement can counteract postural instability despite an unstable centre of gravity, leniency in movement of the arms or stance leg could reduce the reliability and comparability of balance times. Some studies explicitly permitted movement of the legs or arms(Bergland and Wyller, 2004; Heitmann et al., 1989; Sampaio et al., 2013), others ended the trial if there was any movement(Briggs et al., 1989; Choy et al., 2008; Choy et al., 2007; Depasquale and Toscano, 2009; El-Sobkey, 2011; Eto and Miyauchi, 2018; Hasegawa et al., 2019; Hashidate et al., 2011; MacRae et al., 1992; Mulasso et al., 2017; Niam and Wee, 1999), while most studies did not provide details. Similarly, several studies instructed participants to focus their eyes on a head-level target(Ansai et al., 2016; El-Sobkey, 2011; Niam and Wee, 1999); this is hypothesised to improve balance performance as visual concentration can improve proprioceptive input(Wulf and Lewthwaite, 2016; Wulf et al., 2001). There are inconsistent reports of test-retest reliability for the one-legged balance test (intraclass correlation coefficient: 0.56-0.94) (Franchignoni et al., 1998; Kammerlind et al., 2005; Lin et al., 2004a; Wolinsky et al., 2005). Selecting the best result, rather than the average time or a comparison of multiple testing conditions, has been recommended to improve reliability(Ponce-González et al., 2014). However, in the studies synthesised in this review, single trials of balance (n=19) were common, while studies using multiple trials had diverse approaches to selecting a balance score for analysis. Other differences in balance protocols include test duration (e.g. ceiling effects), inclusion of individuals who could not do the test (e.g. zero imputation, exclusion, minimum balance time required for inclusion), testing leg (e.g. left or right, dominant or non-dominant) and cut-points (e.g. <1.02s(Thomas and Lane, 2005) vs. 55.4s(Eto and Miyauchi, 2018)).

A final source of heterogeneity was the ascertainment of falls. Most studies relied on retrospective, self-reported measures; inaccuracies in retrospective recall of falls are common due to poor recollection and interindividual differences in what constitutes a fall (Ganz et al., 2005; Griffin et al., 2019; Sanders et al., 2015). Longer recall periods can further reduce the accuracy of reporting of falls and in addition may contribute to greater residual confounding(Ganz et al., 2005) due to the complexity of factors that accumulate and contribute to subsequent falls(Nowak and Hubbard, 2009). Conversely, if the follow-up period is too short, there may not be sufficient opportunity for a fall event to occur which could lead to associations being underestimated. For example, two studies had a recall or follow-up period of less than 12 months, both Ansai et al. (2016) and Shimada et al. (2011) found no difference in median balance times in fallers and non-fallers over a 3-month recall period.

4.4 Prognostic accuracy and recurrent falls

Traditional analytical techniques such as mean comparison or regression modelling, commonly used in studies identified in this review, do not assess the predictive ability of the one-legged balance test(Grady and Berkowitz, 2011; Ware, 2006). Of the 7 studies that did report the prognostic accuracy of the test, and the 11 studies in which it could be calculated, findings suggest that one-legged balance performance poorly predicts fall outcomes, with low AUCs and higher specificity than sensitivity (Table 3). This indicates that, if used as a screening tool, one-legged balance performance may not adequately identify those at higher risk of falling.

Our synthesis of evidence based on estimates reported in 9 of the 55 included papers suggested that both observational associations and evidence of prognostic accuracy were stronger for recurrent falls than for any fall. This is consistent with previous evidence reporting that individuals who fall one time are more similar to non-fallers than to recurrent or injurious fallers(Delbaere et al., 2010; Lord et al., 1991; Nevitt et al., 1989) and that there are more clearly defined risk factors for recurrent falls(Nevitt et al., 1989; Tinetti and Speechley, 1989). As single falls can commonly occur due to unanticipated environmental hazards, distinct analysis of balance and recurrent falls may better inform overall fall risk. This may also be true for associations between balance and injurious falls, where associations between balance and hospital fall data in a Swedish cohort were robust to adjustment for covariates and follow-up duration(Ek et al., 2019a; Ek et al., 2019b; Welmer et al., 2017). As balance ability may be more consistently associated with recurrent or injurious falls than single falls, allocation of resources to individuals at greater risk of more severe consequences should be considered.

4.5 Strengths and limitations

This systematic review followed a rigorous protocol with two authors independently identifying eligible papers and extracting relevant data on associations. To our knowledge, this is the first systematic review to focus on the one-legged balance test in relation to fall outcomes; as a result, the number of studies identified is much higher compared with reviews that consider multiple balance tests(Gates et al., 2008; Kozinc et al., 2020; Lusardi et al., 2017; Power et al., 2014). For example, a recent systematic review that examined multiple balance tests in relation to fall risk identified 67 studies, only 14 of which examined one-leg balance tests(Kozinc et al., 2020). Another strength of our review is that publication bias was minimised by including all studies that reported on balance-fall associations even if this was not the main study objective.

There are some potential limitations to this review. As only English language articles were included, it is possible that relevant data from non-English languages were missed(Ben Achour Lebib et al., 2006; Hatayama, 2008). Only two meta-analyses could be conducted and we were limited to undertaking narrative syntheses of regression and prognostic estimates due to major heterogeneity between studies in their methods and analytical approaches. Furthermore, publication bias could only be formally assessed using the Egger test and funnel plots for the 17 studies included in the SMD meta-analysis. While comparisons of means need to be interpreted with great caution given the non-parametric distribution of balance times, we decided to report the SMD meta-analysis with caveats as it was the most commonly presented association. Finally, we focused on balance time in relation to fall risk and did not consider other potentially relevant measures of one-legged balance such as postural sway.

4.6 Implications and future steps

Although the one-legged balance test has been recommended as a screening tool for falls in clinical settings outcomes(Kozinc et al., 2020; Michikawa et al., 2009; Nickelston, 2014), we have not found consistent evidence to support this. The results of our review highlight the need for caution and suggest limitations to the use of the one-legged balance test for this purpose in both clinical and research settings. Many studies scored poorly on the Newcastle-Ottawa risk of bias scale (see Appendix B), due to inadequate reporting of balance and fall ascertainment, temporality, low comparability of adjusted estimates and statistical analyses. High-quality longitudinal studies that measure one-legged balance performance before fall reporting periods is crucial to establish temporality of association and minimise the potential impact of reverse causality.

Despite the poor quality of most studies, associations between one-legged balance test and risk of recurrent or injurious falls may be an important avenue of further research. For example, there were robust associations between one-legged balance time and injurious fall risk in the SNAC-K study; whether this is due to the nature of injurious falls or the high quality of cohort data used is not clear. Further investigation of various fall outcomes within the same study sample is necessary to inform translation of this research. If these associations remain, prevention efforts could improve efficiency by targeting those at risk of recurrent or injurious fall outcomes(Peeters et al., 2007).

Few studies examined if associations between balance and falls differed between men and women or at different ages, which is a key consideration when translating findings to clinical settings. One promising avenue for further exploration is the indication that one-legged balance with eyes closed, a more challenging test, may better identify fall risk in younger individuals, while the eyes open test may be a more appropriate test for older adults. This is supported by findings from Cho and Kamen(1998) and El-Sobkey(2011), which reported that balance with eyes closed but not opened was associated with falls in younger adults (mean age: 66.5) and that balance with eyes open but not closed was associated with falls in older adults (mean age 74.5); replication of these analyses in larger, population-representative studies is required. Stratification of the meta-analysis by age (<75, ≥ 75 years) suggested that associations were stronger in younger adults. Although the one-legged balance test is commonly used in those aged 65+, there may be a floor effect at older ages, particularly for those who may be at highest risk of falling. This may partially explain why one-legged balance had poor prognostic accuracy in predicting falls.

Although our review identifies the most common elements of one-legged balance measurement protocols, further work is needed to identify and standardise a protocol for use in research and clinical settings. Factors to consider include number of trials, leg choice, trial duration, continuous timing, body position of arms and raised leg, and criteria for stopping the timed trial. Moving forward, it is equally crucial that all studies who report on one-legged balance tests provide details of the protocol used to better facilitate standardisation and comparison across studies.

A key advantage of the one-legged balance test is its ability to isolate balance ability, in contrast to other measures such as the time-up-and-go, walking speed or chair rise. However, attenuation of estimates after adjustment suggests that other non-balance factors may better explain fall risk. For example, Power et al. (2014) suggested that there was strong evidence that tests that incorporated balance and mobility (e.g. Timed Up and Go, sit to stand or walking speed assessment) could predict falls, with weaker evidence for measures of standing balance and functional reach. While there is utility in examining isolated measures of balance to understand the mechanism of association with falls (Montero-Odasso and Speechley, 2018), a combined risk prediction tool that incorporates balance, mobility and fall history may be preferable. Fall risk screening guidelines have recommended a two-factor approach of fall history and a measure of balance or gait ability (American Geriatrics Society, 2001). If no single test is sufficient to meaningfully predict falls(Gates et al., 2008; Lusardi et al., 2017), further research is needed to create an accurate multifactorial screening tool."

CONCLUSIONS

This systematic review identified 55 papers from 51 study samples that examined the association between one-legged balance performance and fall risk. Study quality was consistently low across papers, limiting our ability to establish any clear conclusions. Despite previous advocacy for the one-legged balance test as a feasible and inexpensive screening tool, we found limited support for this, particularly in studies that temporally distinguished one-legged balance and falls (i.e. longitudinal design). As the global population continues to age, the absence of robust empirical evidence on the association between one-legged balance and falls highlights the need to prioritise high quality studies in this area. Our review highlights crucial gaps in the existing literature that must be addressed to inform translation of balance assessments into effective screening tools to help address the rising prevalence of falls in an ageing population.

Funding: This work was supported by the Canadian Institutes of Health Research (FDSA), the Canadian Centennial Scholarship Fund, the Medical Research Council (MC_UU_00019/1 Theme 1: Cohorts and Data Collection to NSHD; MC_UU_12019/1, MC_UU_12019/2 and MC_UU_12019/4) and the Economic and Social Research Council (ES/K000357/1)

REFERENCES

American Geriatrics Society, B.G.S., and American Academy of Orthopaedic Surgeons Panel on Falls Prevention, 2001. Guideline for the prevention of falls in older persons. J Am Geriatr Soc 49, 664-672.

Andresen, E.M., Wolinsky, F.D., Miller, J.P., Wilson, M.-M.G., Malmstrom, T.K., Miller, D.K., 2006. Cross-sectional and longitudinal risk factors for falls, fear of falling, and falls efficacy in a cohort of middle-aged African Americans. The Gerontologist 46, 249-257.

Ansai, J.H., Aurichio, T.R., Rebelatto, J.R., 2016. Relationship between balance and dual task walking in the very elderly. Geriatrics & gerontology international 16, 89-94.

Arai, T., Fujita, H., Maruya, K., Morita, Y., Asahi, R., Ishibashi, H., 2020. The one-leg portion of the Stand-Up Test predicts fall risk in aged individuals: A prospective cohort study. J Orthop Sci 25, 688-692.

Balduzzi, S., Rücker, G., Schwarzer, G., 2019. How to perform a meta-analysis with R: a practical tutorial. *Evidence-Based Mental Health*, 153-160.

Barry, E., Galvin, R., Keogh, C., Horgan, F., Fahey, T., 2014. Is the timed up and go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta- analysis. *BMC Geriatr* 14.

Beauchet, O., Rossat, A., Bongue, B., Dupre, C., Colvez, A., Fantino, B., Fantino, B., 2010. Change in arm position during one-leg balance test: a predictor of recurrent falls in community-dwelling older adults. J Am Geriatr Soc 58, 1598-1600.

Ben Achour Lebib, S., Missaoui, B., Miri, I., Ben Salah, F.Z., Dziri, C., 2006. Rôle du Neurocom Balance Master dans l'évaluation des troubles de l'équilibre et du risque de chute chez le sujet âgé. A nn Readapt Med Phys 49, 210-217.

Bergland, A., Wyller, T.B., 2004. Risk factors for serious fall related injury in elderly women living at home. Inj Prev 10, 308-313.

Bergquist, R., Weber, M., Schwenk, M., al, e., 2019. Performance-based clinical tests of balance and muscle strength used in young seniors: a systematic literature review. *BMC Geriatr* 19, 9.

Blain, H., Gamon, L., Aliaga, B., Soriteau, L., Raffort, N., Miot, S., Picot, M.C., Bousquet, J., Bernard, P.L., 2021. Self-reported fatigue: A significant risk factor for falling in older women and men. Exp Gerontol 143, 111154.

Blodgett, J.M., Cooper, R., Davis, D.H.J., Kuh, D., Hardy, R., 2020a. Associations between factors across life and balance ability in mid and later life: evidence from a British birth cohort study. Frontiers in Sports and Active Living 2.

Blodgett, J.M., Cooper, R., Davis, D.H.J., Kuh, D., Hardy, R., 2020b. Investigating bidirectional associations between verbal memory and one-legged balance performance in mid and later life. JGMS [submitted].

Blodgett, J.M., Cooper, R., Mills, R., Ventre, J., Hardy, R., 2020c. One-legged balance as a risk factor for falls: a systematic review PROSPERO: CRD42020160413.

Bohannon, R.W., 2006. Single limb stance times - A descriptive meta-analysis of data from individuals at least 60 years of age. Topics in Geriatric Rehabilitation 22, 70-77.

Bongue, B., Dupre, C., Beauchet, O., Rossat, A., Fantino, B., Colvez, A., 2011. A screening tool with five risk factors was developed for fall-risk prediction in community-dwelling elderly. Journal of clinical epidemiology 64, 1152-1160.

Boström, K.J., Dirksen, T., Zentgraf, K., Wagner, H., 2018. The Contribution of Upper Body Movements to Dynamic Balance Regulation during Challenged Locomotion. Frontiers in Human Neuroscience 12.

Boyd, R., Stevens, J.A., 2009. Falls and fear of falling: burden, beliefs and behaviours. Age and Ageing 38, 423–428.

Bridge, P.D., Sawilowsky, S.S., 1999. Increasing physicians' awareness of the impact of statistics on research outcomes: comparative power of the t-test and and Wilcoxon Rank-Sum test in small samples applied research. J Clin Epidemiol 52, 229-235.

Briggs, R.C., Gossman, M.R., Birch, R., Drews, J.E., Shaddeau, S.A., 1989. Balance performance among noninstitutionalized elderly women. Phys Ther 69, 748-756.

Buatois, S., Gueguen, R., Gauchard, G.C., Benetos, A., Perrin, P.P., 2006. Posturography and risk of recurrent falls in healthy non-institutionalized persons aged over 65. Gerontology 52, 345-352.

Buatois, S., Perret-Guillaume, C., Gueguen, R., Miget, P., Vancon, G., Perrin, P., Benetos, A., 2010. A simple clinical scale to stratify risk of recurrent falls in community-dwelling adults aged 65 years and older. Phys Ther 90, 550-560.

Cho, C.Y., Kamen, G., 1998. Detecting balance deficits in frequent fallers using clinical and quantitative evaluation tools. Journal of the American Geriatrics Society 46, 426-430.

Choi, Y.M., Dobson, F., Martin, J., Bennell, K.L., Hinman, R.S., 2014. Interrater and intrarater reliability of common clinical standing balance tests for people with hip osteoarthritis. Phys Ther 94, 696-704.

Choy, N.L., Brauer, S., Nitz, J., 2008. Linking stability to demographics, strength and sensory system function in women over 40 to support pre-emptive preventive intervention. Climacteric 11, 144-154.

Choy, N.L.L., Brauer, S.G., Nitz, J.C., 2007. Timed stance performances reflect differences in age, prevalence of comorbidities, medication use, fall history and activity level: Early screening for balance loss is indicated. Australasian Journal on Ageing 26, 29-34.

Cooper, R., Hardy, R., Aihie Sayer, A., Ben-Shlomo, Y., Birnie, K., Cooper, C., Craig, L., Deary, I.J., Demakakos, P., Gallacher, J., McNeill, G., Martin, R.M., Starr, J.M., Steptoe, A., Kuh, D., team, H.A.s., 2011. Age and gender differences in physical capability levels from mid-life onwards: the harmonisation and meta-analysis of data from eight UK cohort studies. PLoS One 6, e27899.

Crenshaw, J.R., Bernhardt, K.A., Atkinson, E.J., Achenbach, S.J., Khosla, S., Amin, S., Kaufman, K.R., 2020. Posterior single-stepping thresholds are prospectively related to falls in older women. Aging Clin Exp Res 32, 2507-2515.

de Rekeneire, N., Visser, M., Peila, R., Nevitt, M.C., Cauley, J.A., Tylavsky, F.A., Simonsick, E.M., Harris, T.B., 2003. Is a fall just a fall: correlates of falling in healthy older persons. The Health, Aging and Body Composition Study. J Am Geratri Soc 51, 841-846.

Delbaere, K., Close, J.C., Heim, J., Sachdev, P.S., Brodaty, H., Slavin, M.J., Kochan, N.A., Lord, S.R., 2010. A multifactorial approach to understanding fall risk in older people. J Am Geriatr Soc 58, 1679-1685.

Dellinger, A., 2017. Older Adult Falls: Effective Approaches to Prevention. Curr Trauma Rep 3, 118-123.

Depasquale, L., Toscano, L., 2009. The Spring Scale Test: a reliable and valid tool for explaining fall history. Journal of geriatric physical therapy (2001) 32, 159-167.

Ek, S., Rizzuto, D., Calderon-Larranaga, A., Franzen, E., Xu, W., Welmer, A.K., 2019a. Predicting First-Time Injurious Falls in Older Men and Women Living in the Community: Development of the First Injurious Fall Screening Tool. Journal of the American Medical Directors Association 20, 1163-1168.e1163.

Ek, S., Rizzuto, D., Fratiglioni, L., Calderon-Larranaga, A., Johnell, K., Sjoberg, L., Xu, W., Welmer, A.-K., 2019b. Risk factors for injurious falls in older adults: the role of sex and length of follow-up. J Am Geriatr Soc 67, 246-253.

El-Sobkey, S.B., 2011. Balance performance of community-dwelling older people. Saudi medical journal 32, 283-287.

Eto, M., Miyauchi, S., 2018. Relationship between occlusal force and falls among community-dwelling elderly in Japan: a cross-sectional correlative study. BMC geriatrics 18, 111.

Florence, C.S., Bergen, G., Atherly, A., Burns, E., Stevens, J., Drake, C., 2018. Medical Costs of Fatal and Nonfatal Falls in Older Adults. *J Am Geriatr Soc* 66, 693-698.

Franchignoni, F., Tesio, L., Martino, M.T., Ricupero, C., 1998. Reliability of four simple, quantitative tests of balance and mobility in healthy elderly females. Aging Clinical and Experimental Research 10, 26-31.

Ganz, D.A., Bao, Y., Shekelle, P.G., Rubenstein, L.Z., 2007. Will my patient fall? JAMA 297, 77-86.

Ganz, D.A., Higashi, T., Rubenstein, L.Z., 2005. Monitoring falls in cohort studies of community-dwelling older people: effect of the recall interval. J Am Geriatr Soc 53, 2190-2194.

Gates, S., Smith, L.A., Fisher, J.D., Lamb, S.E., 2008. Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. The Journal of Rehabilitation Research and Development 45, 1105.

Gerdhem, P., Ringsberg, K.A.M., Akesson, K., Obrant, K.J., 2005. Clinical history and biologic age predicted falls better than objective functional tests. Journal of clinical epidemiology 58, 226-232.

Grady, D., Berkowitz, S.A., 2011. Why is a good clinical prediction rule so hard to find? Arch Intern Med 171, 1701-1702.

Griffin, J., Lall, R., Bruce, J., Withers, E., Finnegan, S., Lamb, S.E., Underwood, M., Martin, F., Yardley, L., Skelton, D., Willett, K., Eldridge, S., Slowther, A.-M., Duggan, S., Hennings, S., Mant, R., Rai, R., Turner, C., Andrews, A., Fearn, R., Walker, N., Potter, R., Hulme, C., Bojke, C., Longo, R., Westacott, K., Ralhan, S., Sheridan, R., Treml, J., Riglin, J., Gordjin, H., Dutta, R., Burns, J., Shaw, F., Davison, J., Willis, A., Muthiah, C., Adjei, H., 2019. Comparison of alternative falls data collection methods in the Prevention of Falls Injury Trial (PreFIT). Journal of Clinical Epidemiology 106, 32-40.

Hasegawa, Y., Horii, N., Sakuramoto-Sadakane, A., Nagai, K., Ono, T., Sawada, T., Shinmura, K., Kishimoto, H., 2019. Is a history of falling related to oral function? A cross-sectional survey of elderly subjects in rural Japan. International Journal of Environmental Research and Public Health 16.

Hashidate, H., Shimada, H., Shiomi, T., Sasamoto, N., 2011. Usefulness of the subjective risk rating of specific tasks for falls in frail elderly people. Journal of Physical Therapy Science 23, 519-524.

Hatayama, T.N., M Une, H Yoshitake, Y Kimura, Y Momose, Y Kai, Y Suwa, M Kumagai, S, 2008. The association between incidence of falls and physical fitness in community- dwelling elderly. Japanese Journal of physical fitness and sports medicine 57, 503-510.

Heitmann, D.K., Gossman, M.R., Shaddeau, S.A., Jackson, J.R., D.K, H., M.R, G., S.A, S., Dk, H., Mr, G., Sa, S., Jr, J., 1989. Balance performance and step width in noninstitutionalized, elderly, female fallers and nonfallers. Physical Therapy 69, 923-931.

Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ 327, 557-560.

Hopewell, S., Adedire, O., Copsey, B.J., Boniface, G.J., Sherrington, C., Clemson, L., Close, J.C., Lamb, S.E., 2018. Multifactorial and multiple component interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev 7, Cd012221.

Ikegami, S., Takahashi, J., Uehara, M., Tokida, R., Nishimura, H., Sakai, A., Kato, H., 2019. Physical performance reflects cognitive function, fall risk, and quality of life in community-dwelling older people. Scientific Reports 9, 12242-12242.
Jalali, M.M., Gerami, H., Heidarzadeh, A., Soleimani, R., 2015. Balance performance in older adults and its relationship with falling. Aging clinical and experimental research 27, 287-296.
James, S.L., Lucchesi, L.R., Bisignano, C., 2020. The global burden of falls: global, regional and national estimates of morbidity and mortality from the Global Burden of Disease Study 2017 *Inj Prev* 26, 3-11.
Jonsson, E., Seiger, A., Hirschfeld, H., 2004. One-leg stance in healthy young and elderly adults: a measure of postural steadiness? Clin Biomech (Bristol, Avon) 19, 688-694.
Kammerlind, A.S., Bergquist Larsson, P., Ledin, T., Skargren, E., 2005. Reliability of clinical balance tests and subjective ratings in dizziness and disequilibrium. Advances in Physiotherapy 7, 96-107.
Kozinc, Ž., Löfler, S., Hofer, C., Carraro, U., Šarabon, N., 2020. Diagnostic balance tests for assessing risk of falls and distinguishing older adult fallers and non-fallers: a systematic review with meta-analysis. *Diagnostics (Basel)* 10, 667.
Kwan, M.M.-S., Lin, S.-I., Chen, C.-H., Close, J.C., Lord, S.R., 2011. Minimal chair height standing ability is independently associated with falls in Taiwanese older people. Archives of physical medicine and rehabilitation 92, 1080-1085.
Lakens, D., 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in Psychology 4, 863.
Li, F., He, H., 2018. Assessing the Accuracy of Diagnostic Tests. *Shanghai archives of psychiatry* 30, 207–212.
Lim, Y., Kim, K., Ko, S.-H., Cho, K., Jang, E.-H., Lee, S.-H., Lim, D.J., Baek, K.H., Ha, H.-S., Park, M.S., Yim, H.-W., Lee, W.-C., Yoon, K.-H., Son, H.Y., Oh, K.W., Kang, M.-I., 2016. Gender- and age-group-specific associations between physical performance and bone mineral density, falls, and osteoporotic fractures in Koreans: the Chungju Metabolic Disease Cohort study. Journal of bone and mineral metabolism 34, 336-346.
Lima, C.A., Ricci, N.A., Nogueira, E.C., Perracini, M.R., 2018. The Berg Balance Scale as a clinical screening tool to predict fall risk in older adults: a systematic review. Physiotherapy 104, 383-394.
Lin, M.-R., Hwang, H.-F., Hu, M.-H., Wu, H.-D.I., Wang, Y.-W., Huang, F.-C., 2004a. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. Journal of the American Geriatrics Society 52, 1343-1348.
Lin, M.R., Hwang, H.F., Hu, M.H., Wu, H.D., Wang, Y.W., Huang, F.C., 2004b. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. J Am Geriatr Soc 52, 1343-1348.
Lord, S.R., Clark, R.D., Webster, I.W., 1991. Physiological factors associated with falls in an elderly population. J Am Geriatr Soc 39, 1194-1200.
Lusardi, M.M., Fritz, S., Middleton, A., Allison, L., Wingood, M., Phillips, E., Criss, M., Verma, S., Osborne, J., Chui, K.K., 2017. Determining risk of falls in community dwelling older adults: a systematic review and meta-analysis using posttest probability. J Geriatr Phys Ther 40, 1-36.
Ma, J., Liu, W., Hunter, A., *et al*, 2008. Performing meta-analysis with incomplete statistical information in clinical trials. *BMC Med Res Methodol* 8.

MacRae, P.G., Lacourse, M., Moldavon, R., 1992. Physical performance measures that predict faller status in community- dwelling older adults. Journal of Orthopaedic and Sports Physical Therapy 16, 123-128.
Mahoney, J.R., Cotton, K., Verghese, J., 2019. Multisensory Integration Predicts Balance and Falls in Older Adults. The journals of gerontology Series A, Biological sciences and medical sciences. 74, 1429-1435.
Mancini, M., Horak, F.B., 2010. The relevance of clinical balance assessment tools to differentiate balance deficits. *Eur J Phys Rehabil Med* 46, 239-248.
McGrath, S., Sohn, H., Steele, R., Benedetti, A., 2020. Meta-analysis of the difference of medians. *Biometrical Journal* 62, 69 – 98.
Michikawa, T., Nishiwaki, Y., Takebayashi, T., Toyama, Y., 2009. One-leg standing test for elderly populations. J Orthop Sci 14, 675-685.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The, P.G., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6.
Montero-Odasso, M., Speechley, M., 2018. Falls in cognitively impaired older adults: implications for risk sssessment and prevention. J Am Geriatr Soc 66, 367-375.
Moore, M., Barker, K., 2017. The validity and reliability of the four square step test in different adult populations: a systematic review. Syst Rev 6, 187.
Moreira, M.N., Bilton, T.L., Dias, R.C., Ferriolli, E., Perracini, M.R., 2017. What are the main physical functioning factors associated with falls among older people with different perceived fall risk? Physiother Res Int 22, e1664.
Morioka, S., Fukumoto, T., Hiyamizu, M., Matsuo, A., Takebayashi, H., Miyamoto, K., 2012. Changes in the equilibrium of standing on one leg at various life stages. *Curr Gerontol Geriatr Res* 2012, 516283.
Muir, S.W., Berg, K., Chesworth, B., Klar, N., Speechley, M., 2010. Balance impairment as a risk factor for falls in community-dwelling older adults who are high functioning: a prospective study. Physical Therapy 90, 338-347.
Mulasso, A., Roppolo, M., Gobbens, R.J., Rabaglietti, E., 2017. Mobility, balance and frailty in community-dwelling older adults: What is the best 1-year predictor of falls? Geriatrics & Gerontology International 17, 1463-1469.
Nahm, F.S., 2016. Nonparametric statistical tests for the continuous data: the basic concept and the practical use. Korean J Anesthesiol 69, 8-14.
Nevitt, M.C., Cummings, S.R., Kidd, S., Black, D., 1989. Risk factors for recurrent nonsyncopal falls. A prospective study. JAMA 261, 2663-2668.
Niam, S., Wee, S.K., 1999. Relationship of unipedal stance time, age and falls in healthy Asian elderly. Physiotherapy Singapore 2, 124-127.
NICE guideline, 2013. Falls in older people: assessing risk and prevention, in: [CG161], C.g. (Ed.). Accessed July 2020 at https://www.nice.org.uk/guidance/CG161/chapter/introduction.
Nickelston, P., 2014. Why You Should Include the Single-Leg Stance Test in Every Patient Assessment. Dynamic Chiropractic 32, 8-14.
No author, 1987. The prevention of falls in later life. A report of the Kellogg International Work Group on the Prevention of Falls by the Elderly. *Dan Med Bull* 34, 1-24.
Nowak, A., Hubbard, R.E., 2009. Falls and frailty: lessons from complex systems. Journal of the Royal Society of Medicine 102, 98-102.
Okubo, Y., Schoene, D., Caetano, M.J., Pliner, E.M., Osuka, Y., Toson, B., Lord, S.R., 2021. Stepping impairment and falls in older adults: A systematic review and meta-analysis of volitional and reactive step tests. Ageing Research Reviews 66, 101238.
Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A., 2016. Rayyan—a web and mobile app for systematic reviews. *Syst Rev* 5.

Overstall, P.W., Exton-Smith, A.N., Imms, F.J., Johnson, A.L., 1977. Falls in the elderly related to postural imbalance. BMJ, 261-264.

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., Thomas, R., 2008. Understanding and using sensitivity, specificity and predictive values. Indian journal of ophthalmology 56, 45-50.

Park, K., Yang, M., Yoo, T.-g., Kim, S.-h., 2020. One-Leg Standing and Y-Balance Test Performance in Elderly Fallers and Nonfallers. Topics in Geriatric Rehabilitation 36, 92-96.

Peeters, G., Cooper, R., Tooth, L., van Schoor, N.M., Kenny, R.A., 2019. A comprehensive assessment of risk factors for falls in middle-aged adults: co-ordinated analyses of cohort studies in four countries. Osteoporos Int.

Peeters, G., van Schoor, N.M., Cooper, R., Tooth, L., Kenny, R.A., 2018. Should prevention of falls start earlier? Co-ordinated analyses of harmonised data on falls in middle-aged adults across four population-based cohort studies. PLoS One 13, e0201989.

Peeters, G.M., de Vries, O.J., Elders, P.J., Pluijm, S.M., Bouter, L.M., Lips, P., 2007. Prevention of fall incidents in patients with a high risk of falling: design of a randomised controlled trial with an economic evaluation of the effect of multidisciplinary transmural care. BMC Geriatr 7, 15.

Ponce-González, J.G., Sanchis-Moysi, J., González-Henriquez, J.J., Arteaga-Ortiz, R., Calbet, J.A., Dorado, C., 2014. A reliable unipedal stance test for the assessment of balance using a force platform. J Sports Med Phys Fitness 54, 108-117.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., 2006. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews, A Product from the ESRC Methods Programme. Version 1.

Porto, J.M., Iosimuta, N.C.R., Freire Júnior, R.C., Braghin, R.M.B., Leitner, É., Freitas, L.G., de Abreu, D.C.C., 2020. Risk factors for future falls among community-dwelling older adults without a fall in the previous year: A prospective one-year longitudinal study. Arch Gerontol Geriatr 91, 104161.

Power, V., Van De Ven, P., Nelson, J., Clifford, A.M., 2014. Predicting falls in community-dwelling older adults: a systematic review of task performance-based assessment tools. Physiotherapy Practice 35, 3-15.

Rohatgi, A., 2020. WebPlotDigitalizer: HTML5 based online tool to extract numerical data from plot images. Version 4.3., p. Accessed http:// arohatgi.info/WebPlotDigitizer/app/ in June 2020.

Rosa, M.V., Perracini, M.R., Ricci, N.A., 2019. Usefulness, assessment and normative data of the Functional Reach Test in older adults: A systematic review and meta-analysis. Arch Gerontol Geriatr 81, 149-170.

Rossat, A., Fantino, B., Bongue, B., Colvez, A., Nitenberg, C., Annweiler, C., Beauchet, O., 2010. Association between benzodiazepines and recurrent falls: A cross-sectional elderly population-based study. Journal of Nutrition, Health and Aging, 1-6.

Sampaio, R.A.C., Sampaio, P.Y.S., Yamada, M., Ogita, M., Matsudo, S.M.M., Raso, V., Tsuboyama, T., Arai, H., 2013. Factors associated with falls in active older adults in Japan and Brazil. Journal of Clinical Gerontology and Geriatrics 4, 89-92.

Sanders, K.M., Stuart, A.L., Scott, D., Kotowicz, M.A., Nicholson, G.C., 2015. Validity of 12-Month Falls Recall in Community-Dwelling Older Women Participating in a Clinical Trial. Int J Endocrinol 2015, 210527.

Scholz, J.P., Park, E., Jeka, J.J., Schöner, G., Kiemel, T., 2012. How visual information links to multijoint coordination during quiet standing. Exp Brain Res 222, 229-239.

Sherrington, C., Fairhall, N.J., Wallbank, G.K., Tiedemann, A., Michaleff, Z.A., Howard, K., Clemson, L., Hopewell, S., Lamb, S.E., 2019. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 1, Cd012424.

Shimada, H., Suzukawa, M., Tiedemann, A., Kobayashi, K., Yoshida, H., Suzuki, T., 2009. Which neuromuscular or cognitive test is the optimal screening tool to predict falls in frail community-dwelling older people? Gerontology 55, 532-538.

Shimada, H., Tiedemann, A., Lord, S.R., Suzukawa, M., Makizako, H., Kobayashi, K., Suzuki, T., 2011. Physical factors underlying the association between lower walking performance and falls in older people: a structural equation model. Archives of gerontology and geriatrics 53, 131-134.

Shin, K.R., Kang, Y., Jung, D., Kim, M., Lee, E., 2012. A Comparative Study on Physical Function Test between Faller Group and Nonfaller Group among Community-Dwelling Elderly. Asian nursing research 6, 42-48.

Shinohara, T., Saida, K., Miyata, K., 2020. Ability of the Brief-Balance Evaluation Systems Test to evaluate balance deficits in community-dwelling older adults: a cross-sectional study. Physiother Theory Pract, 1-8.

Soriano, T.A., DeCherrie, L.V., Thomas, D.C., 2007. Falls in the community-dwelling older adult: a review for primary-care providers. Clin Interv Aging 2, 545-554.

Springer, B.A., Marin, R., Cyhan, T., Roberts, H., Gill, N.W., 2007. Normative values for the unipedal stance test with eyes open and closed. Journal of Geriatric Physical Therapy 30, 8-15.

Stalenhoef, P.A., Diederiks, J.P., Knottnerus, J.A., Kester, A.D., Crebolder, H.F., 2002. A risk model for the prediction of recurrent falls in community-dwelling elderly: a prospective cohort study. J Clin Epidemiol 55, 1088-1094.

Sterne, J.A., Egger, M., 2005. Regression Methods to Detect Publication and Other Bias in Meta-Analysis, Publication Bias in Meta-Analysis, pp. 99-110.

Swanenburg, J., Nevzati, A., Mittaz Hager, A.G., de Bruin, E.D., Klipstein, A., 2013. The maximal width of the base of support (BSW): clinical applicability and reliability of a preferred-standing test for measuring the risk of falling. Archives of gerontology and geriatrics 57, 204-210.

Thomas, J.I., Lane, J.V., 2005. A pilot study to explore the predictive validity of 4 measures of falls risk in frail elderly patients. Arch Phys Med Rehabil 86, 1636-1640.

Tinetti, M.E., Speechley, M., 1989. Prevention of falls among the elderly. N Engl J Med 320, 1055-1059.

Tinetti, M.E., Speechley, M., Ginter, S.F., 1988. Risk factors for falls amongst elderly persons living in the community. New England Journal of Medicine 319, 1701-1707.

Toulotte, C., Thevenon, A., Watelain, E., Fabre, C., 2006. Identification of healthy elderly fallers and non-fallers by gait analysis under dual-task conditions. Clinical rehabilitation 20, 269-276.

Ungar, A., Rafanelli, M., Iacomelli, I., Brunetti, M.A., Ceccofiglio, A., Tesi, F., Marchionni, N., 2013. Fall prevention in the elderly. Clin Cases Miner Bone Metab 10, 91-95.

Vellas, B.J., Wayne, S.J., Garry, P.J., Baumgartner, R.N., 1998. A two-year longitudinal study of falls in 482 community-dwelling elderly adults. The journals of gerontology. Series A, Biological sciences and medical sciences 53, M264-274.

Vellas, B.J., Wayne, S.J., Romero, L., Baumgartner, R.N., Rubenstein, L.Z., Garry, P.J., 1997. One-leg balance is an important predictor of injurious falls in older persons. J Am Geriatr Soc 45, 735-738.

Vereeck, L., Wuyts, F., Truijen, S., Van de Heyning, P., 2008. Clinical assessment of balance: normative data, and gender and age effects. Int J Audiol 47, 67-75.

Vickers, A.J., 2005. Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data. B*MC Med Res Methodol* 5.

Ware, J.H., 2006. The limitations of risk factors as prognostic tools. N Engl J Med 355, 2615-2617.

Welmer, A.K., Rizzuto, D., Laukka, E.J., Johnell, K., Fratiglioni, L., 2017. Cognitive and physical function in relation to the risk of injurious falls in older adults: a population-based study. J Gerontol A Biol Sci Med Sci 72, 669-675.

WHO, 2007. WHO Global Report on Falls Prevention in Older Age.

WHO, 2021. Health Topics Fact Sheet: Falls.

Wolinsky, F.D., Miller, D.K., Andresen, E.M., Malmstrom, T.K., Miller, J.P., 2005. Reproducibility of physical performance and physiologic assessments. J Aging Health 17, 111-124.

Wulf, G., Lewthwaite, R., 2016. Optimizing performance through intrinsic motivation and attention for learning: the optimal theory of motor learning. *Psychon. Bull. Rev.* 23, 1382–1414.

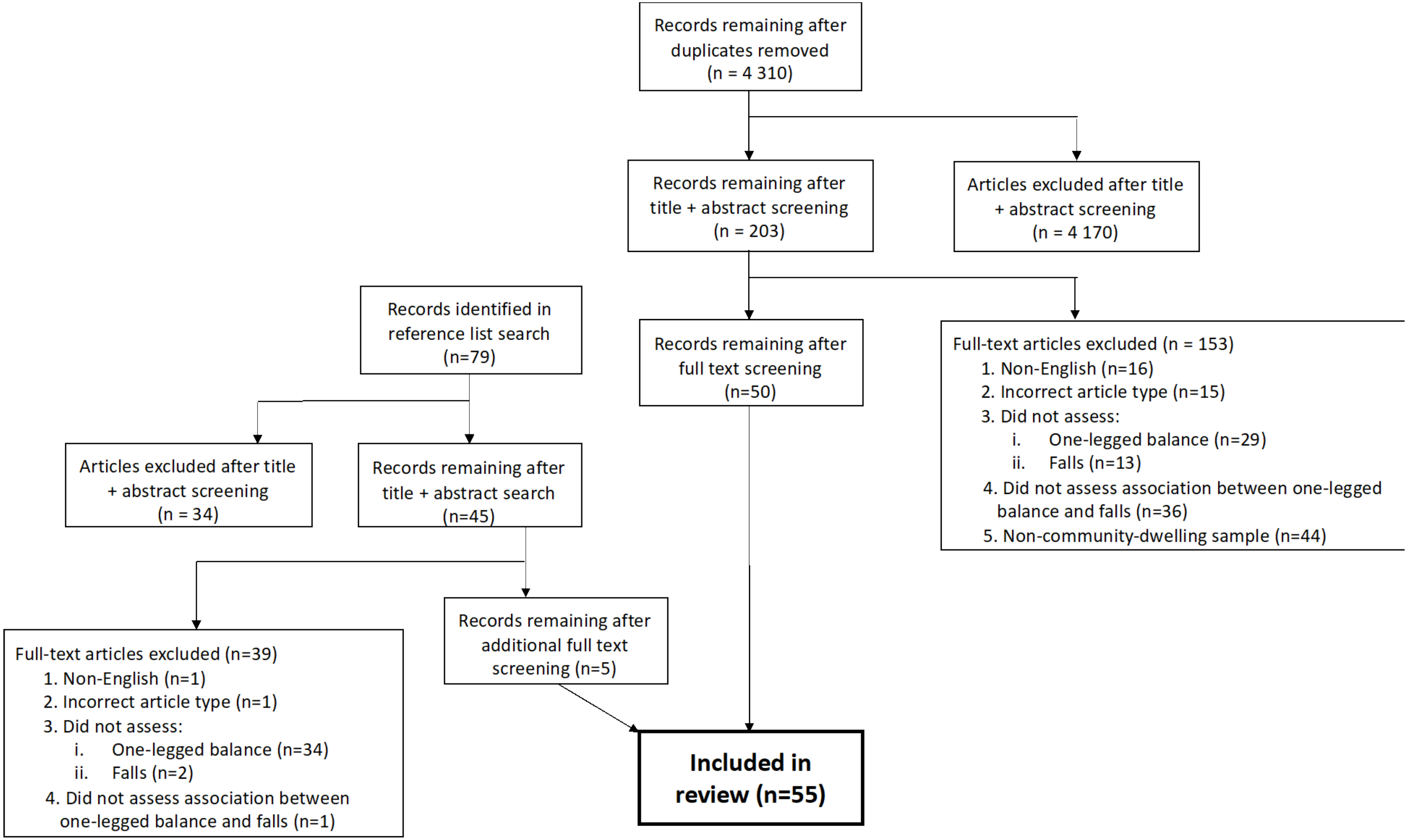
Wulf, G., McNevin, N.H., Shea, C.H., 2001. The automaticity of complex motor skill learning as a function of attentional focus. *Q. J. Exp. Psychol* 54A, 1143–1154.

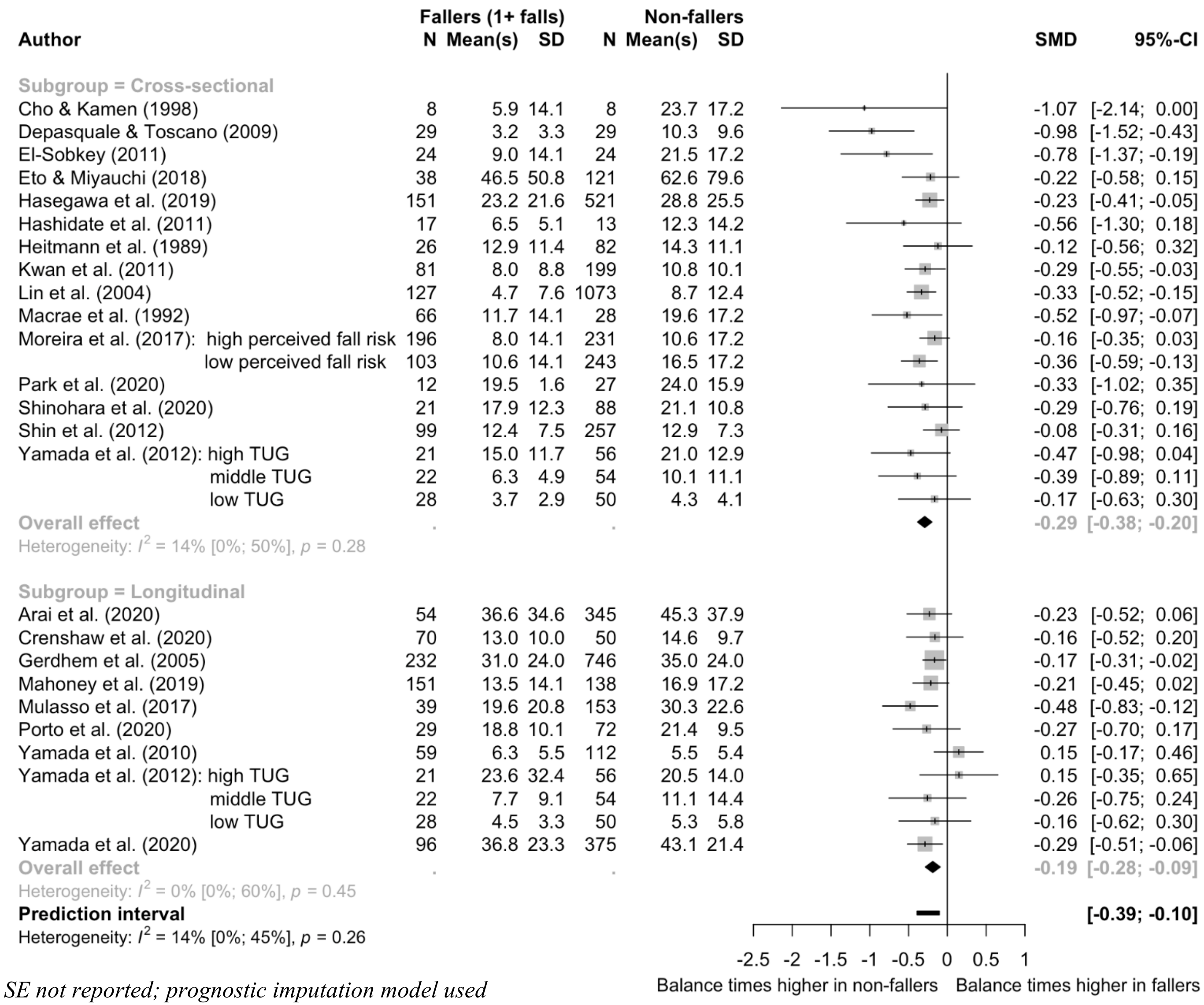
Yamada, M., Ichihashi, N., 2010. Predicting the probability of falls in community-dwelling elderly individuals using the trail-walking test. Environmental health and preventive medicine 15, 386-391.

Yamada, M., Uemura, K., Mori, S., Nagai, K., Uehara, T., Arai, H., Aoyama, T., 2012. Faster decline of physical performance in older adults with higher levels of baseline locomotive function. Geriatrics & gerontology international 12, 238-246.

Yamada, T., Yamato, Y., Hasegawa, T., Yoshida, G., Yasuda, T., Banno, T., Arima, H., Oe, S., Ushirozako, H., Ide, K., Watanabe, Y., Matsuyama, Y., 2020. Influence of the Sagittal Vertical Axis on the Risk of Falls in Community-Dwelling Elderly People: A Retrospective Longitudinal Study. Spine Surg Relat Res 4, 237-241.

Figure 1. PRISMA flow chart outlining identification of eligible studies

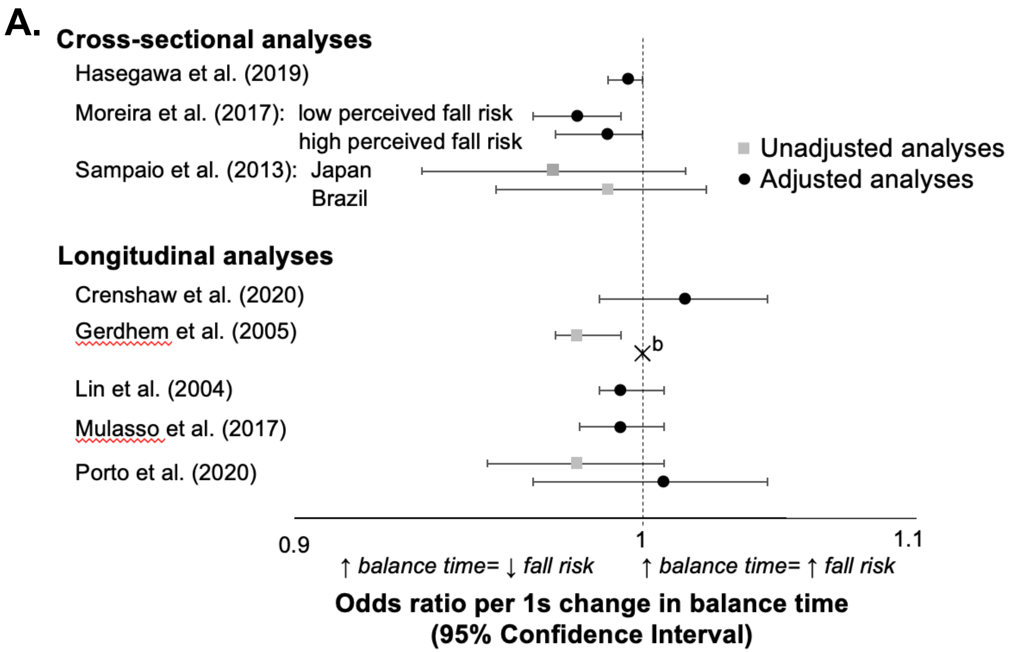


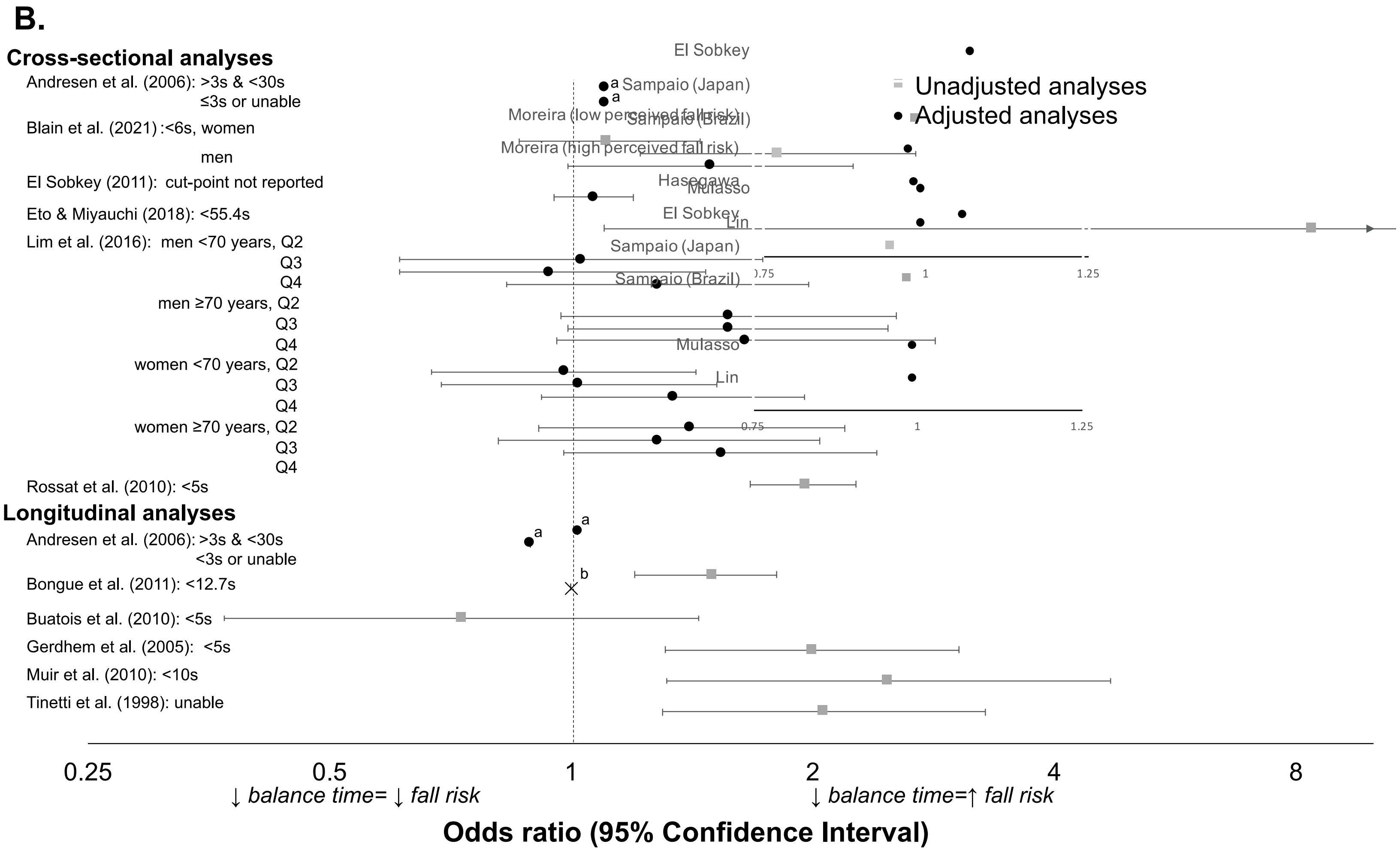


[a] *SD or SE not reported; prognostic imputation model used*

TUG Timed-up-and-go

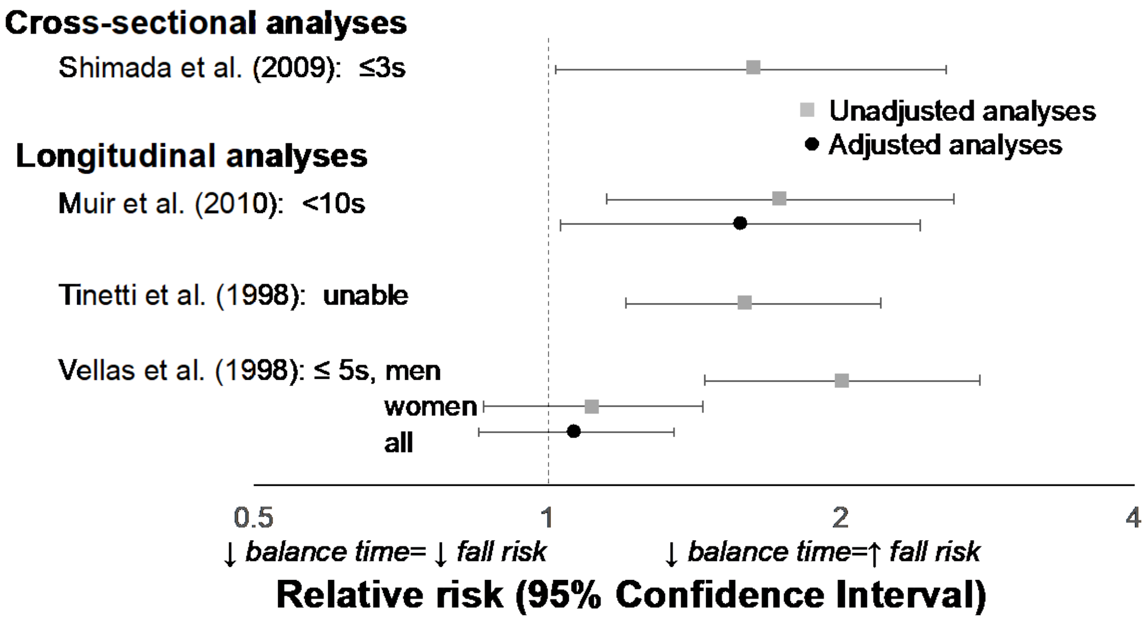
Figure 2. Forest plot showing standardised mean difference in one-legged balance times between fallers (1+ fall) and non-fallers.


A.
Cross-sectional analyses
Hasegawa et al. (2019)
Moreira et al. (2017): low perceived fall risk
high perceived fall risk
Sampaio et al. (2013): Japan
Brazil
Longitudinal analyses
Crenshaw et al. (2020)
Gerdhem et al. (2005)
b
Lin et al. (2004)
Mulasso et al. (2017)
Porto et al. (2020)
Unadjusted analyses
Adjusted analyses
0.9
1
1.1
↑ balance time= ↓ fall risk
↑ balance time= ↑ fall risk
Odds ratio per 1s change in balance time
(95% Confidence Interval)



B.
Cross-sectional analyses
Andresen et al. (2006): >3s & <30s
≤3s or unable
Blain et al. (2021) :<6s, women
men
El Sobkey (2011): cut-point not reported
Eto & Miyauchi (2018): <55.4s
Lim et al. (2016): men <70 years, Q2
Q3
Q4
men ≥70 years, Q2
Q3
Q4
women <70 years, Q2
Q3
Q4
women ≥70 years, Q2
Q3
Q4
Rossat et al. (2010): <5s
Longitudinal analyses
Andresen et al. (2006): >3s & <30s
<3s or unable
Bongue et al. (2011): <12.7s
Buatois et al. (2010): <5s
Gerdhem et al. (2005): <5s
Muir et al. (2010): <10s
Tinetti et al. (1998): unable
Unadjusted analyses
Adjusted analyses
0.25
0.5
1
2
4
8
↓ balance time= ↓ fall risk
↓ balance time=↑ fall risk
Odds ratio (95% Confidence Interval)

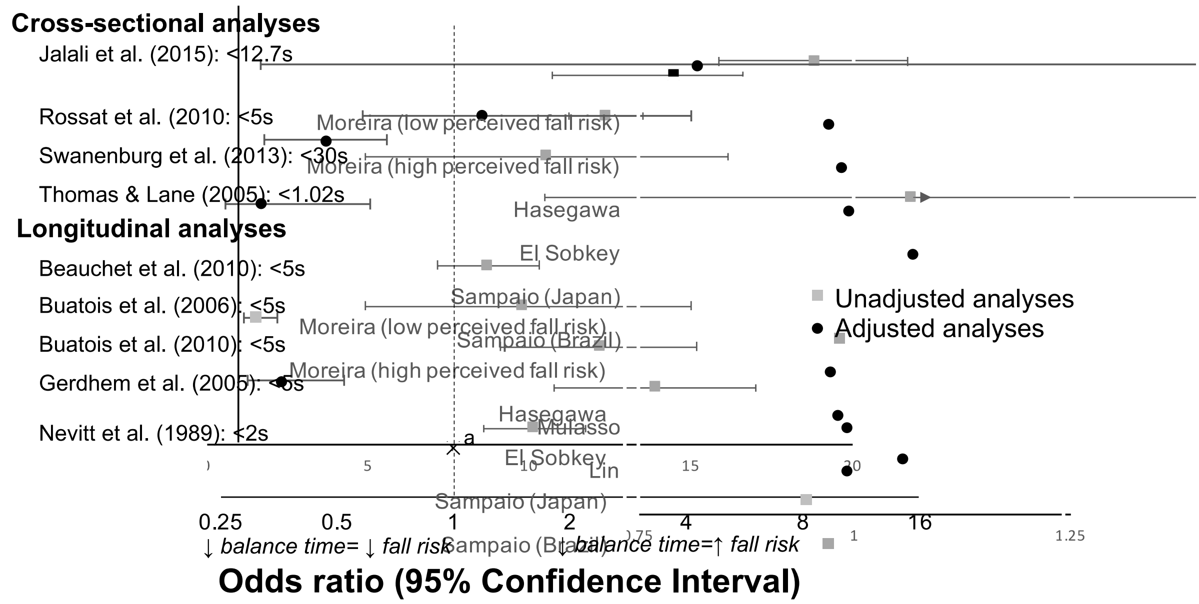
C.



[a] Andresen: no 95% CI provided; estimates not significant.
[b] Bongue and Gerdhem: adjusted model not reported and not significant;

Reference categories: Andresen 30s; Blain 6-10s; Eto 55.4-120s; Lim highest quartile in age and sex-specific groups (Q1); Rossat 5s; Bongue, 12.7-60s; Buatois 5s; Gerdhem 5-30s; Muir 10s; Tinetti able to stand on one leg; Shimada 3-120s; Vellas 5s;

Figure 3. Risk of any fall (1+): A. Odds ratio per 1s increase in balance time; B. Odds ratio per low balance cut-point; C. Relative risk per low balance cut-point (see study details in Appendix F)



[a] *adjusted model not reported and not significant*

Figure 4. Odds ratios of recurrent falls (2+ falls) in those with low one-legged balance times compared to single or non-fallers (0-1 falls)

Table 1. Characteristics of included papers (n=55; listed in alphabetic order of first author surnames)

First author, year **Study country** (and study name where applicable)	**Analytical sample size** (% women) **Age** (mean ±SD, min-max) *nr = not recorded*	**One-legged balance:** **assessed time** (continuous range or categorical cut-points) **analysed time** (if different) *nr = not recorded*	**Eyes open or closed**	**Fall temporality:** **Retrospective** or **Prospective** (recall or follow-up period)	**Fall ascertainment**	**Fall outcomes analysed**	**Temporality of analyses** (cross-sectional or longitudinal) a	**Quality score (0-7)**
Andresen, 2006 USA African American Health cohort study	998 (58.2%) 56.8 ± 4.4, *nr*	Continuous (0-30s) Analysed categorical (unable or ≤3s; 3-29s; 30s)	Not stated	Retrospective (2 years) Prospective (2 years)	Self-reported in interview Received annual phone calls	0 falls; 1+ falls 0 or non-injurious falls; 1+ injurious falls 0 falls; 1+ falls	Cross-sectional Cross-sectional Longitudinal	5
Ansai, 2016 Brazil	67 (67.2%) *nr* ± *nr,* 80+	Continuous (0-30s)	Open	Retrospective (3 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	3
Arai, 2020 Japan	399 (52.4%) 71.7 ± 4.2, 65-79	Continuous (0-120s)	Open	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Longitudinal	5
Beauchet, 2010 France	1759 (51.0%) 70.7 ± 4.6, 65-95	Binary (<5; 5s) Moved arms (yes/no)	Not stated	Prospective (12 months)	Received monthly phone calls	0 falls; 1 fall; 2+ falls	Longitudinal	4
Bergland, 2004 Norway	307 (100%) 80.8 ± *nr,* 75-93	Continuous (0-*nr*) Analysed binary (median cut-off; not stated)	Open	Prospective (12 months)	Submitted falls diary every 3 months	0 or non-injurious falls; 1+ injurious falls	Longitudinal	7
Blain, 2021	1471 (67.0%)	Continuous	Not stated	Retrospective	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	4

France	72.4 ± 5.1, 65-*nr*	(0-10s) Analysed binary (<6.5; ≥ 6.5s)		(12 months)				
Bongue, 2011 France	1759 (51%) 70.7 ± 4.6, 65-95	Continuous (0-60s) Analysed binary (Dominant leg: <12.7; ≥12.7s Non-dominant leg: <7.6; ≥7.6s)	Open	Prospective (12 months)	Received monthly phone calls	0 falls; 1+ falls	Longitudinal	7
Briggs, 1989 USA	71 (100%) 72.3 ± 7.0, 60-86	Continuous (0-45s)	Open + Closed	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional	3
Buatois, 2006 France	189 (43.7%) 70.0 ± 4.0, 65+	Binary (<5; 5s)	Open	Prospective (16 months)	Responded to questionnaire every 4 months	0 falls; 1 fall; 2+ falls	Longitudinal	4
Buatois, 2010 France	1618 (49.3%) 70.3 ± 4.5, 65+	Binary (<5; 5s)	Not stated	Retrospective (mean 25±5 months)	Self-reported questionnaire	0-1 falls; 2+ falls	Longitudinal	4
Cho, 1998 Country not stated	16 (75%) 74.5 ± *nr,* 65-87	Continuous (0-30s)	Open + Closed	Retrospective (2 years)	Clinician referral of recurrent fallers	0-1 falls; 2+ falls	Cross-sectional	1
Choy, 2007 Australia	456 (100%) *nr* ± *nr,* 20-80	Binary (<10; 10s) Analysed categorical (stable, unsteady, unstable)	Open + Closed	Retrospective (12 months)	Self-reported questionnaire	Continuous # of falls	Cross-sectional	1
Choy, 2008 Australia	254 (100%) *nr* ± *nr,* 40-80	Binary (<10; 10s) Analysed categorical (stable=3 successful trials; unsteady =1-2 successful trials;	Open	Retrospective (12 months)	Self-reported questionnaire	Continuous # of falls	Cross-sectional	1

		unstable=0 successful trials. Where 10s = success)						
Crenshaw, 2020 USA	120 (100%) 77.1 ± 7.5, 65-*nr*	Continuous (0-30s)	Not stated	Prospective (12 months)	Complete biweekly falls questionnaires; received reminder letters and phone calls if questionnaires were missing for a month	0 falls; 1+ falls	Longitudinal	5
Depasquale, 2009 USA	58 (67.2%) 80.8 ± 6.7, 65-94	Continuous (0-30s)	Open	Retrospective (2 years)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional	2
de Rekeneire, 2003 USA Health, Aging & Body Composition Study	3050 (51.5%) *nr* ± *nr,* 70-79	Continuous (0-30s) Analysed categorical (0, 1, 2)	Not stated	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	1
Delbaere, 2006 Australia Sydney Memory and Ageing Study	494 (54%) 77.9 ± 4.1, 70-90	Continuous (0-10s)	Not stated	Prospective (12 months)	Submitted monthly falls diaries	0-1 non-injurious falls; 1 injurious or 2+ falls	Longitudinal	5
Ek, 2019 Sweden SNAC-K	2808 (62.3%) 73 ± 10.3, 60+	Continuous (0-60s) Analysed binary (<5; ≥5s)	Open	Prospective (5 years)	ICD-10 codes via linked health records	0 or non-injurious falls; 1+ injurious falls	Longitudinal	7
Ek, 2019 Sweden SNAC-K	3112 (63.7%) 73.9 ± 10.6, 60+	Continuous (0-60s) Analysed binary (<5; ≥5s)	Open	Prospective (4, 10 years)	ICD-10 codes via linked health records	0 or non-injurious falls; 1+ injurious falls	Longitudinal	7
El Sobkey, 2011 Kingdom of Saudi Arabia	48 (60.4%) 66.5 ± 6.3, 60-85	Continuous (0-45s)	Open + Closed	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls Continuous # of falls	Cross-sectional Cross-sectional	2
Eto, 2018 Japan	159 (64.8%) 74.3 ± 6.3, 65+	Continuous (0-120s)	Open	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	3

Gerdhem, 2005 Sweden	984 (100%) 75 ± *0,* 75-75	Continuous (0-30s) Sum of 4 conditions analysed (0-120s)	Open + Closed	Retrospective (1.01±0.05 years)	Self-reported questionnaire	0 falls; 1+ falls 1 fall; 2+ falls	Longitudinal Longitudinal	6
Hasegawa, 2019 Japan Frail Elderly in the Tamba Sasayama-Area study	672 (66.8%) 72.8 ± 5.9, 65+	Continuous (0-*nr*)	Open	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	4
Hashidate, 2011 Japan	30 (50%) *nr* ± *nr,* 65+	Continuous (0-*nr*)	Open	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional	1
Heitmann, 1989 USA	110 (100%) 73.6 ± 7.2, 60-89	Continuous (0-30s)	Open + Closed	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	2
Ikegami, 2019 Japan Obuse study cohort	*412 (50.7%)* *nr* ± *nr,* 50-89	Continuous (0-60s) Analysed per 1SD	Not stated	Retrospective (12 months)	Self-reported in interview	Continuous # of falls	Cross-sectional	4
Jalali, 2015 Iran	448 (46.7%) 73.8 ± 6.3, 65+	Continuous (0-*nr*) Analysed binary (≤12.7; >12.7s)	Open	Retrospective (12 months)	Self-reported in interview	0-1 falls; 2+ falls	Cross-sectional	4
Kwan, 2011 Taiwan	280 (42.9%) 74.9 ± 6.4, 65-91	Continuous (0-30s)	Open	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	3
Lim, 2016 South Korea Chungju Metabolic Disease Cohort study	5368 (55.8%) 67.7 ± 4.9, 40+	Continuous (0-30s) Analysed categorical (gender & age-specific quartiles)	Not stated	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	3
Lin, 2004 Taiwan	1200 (41%) 73.4 ± *nr,* 65+	Continuous (0-*nr*)	Open	Retrospective (12 months) Prospective	Self-reported in interview	0 falls; 1+ falls 0 falls; 1+ falls	Cross-sectional Longitudinal	6

				(12 months)	Reported each fall by postcard & received phone call every 3 months			
Macrae, 1992 USA	94 (69.1%) 73.2 ± 0.8, 60-89	Continuous (0-30s)	Open	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional	3
Mahoney, 2019 USA Central Control of Mobility in Aging study	289 (53%) 76.7 ± 6.4, 65-93	Continuous (0-30s)	Not stated	Prospective (24±17 months)	Received phone call every 2-3 months	0 falls; 1+ falls	Longitudinal	6
Moreira, 2005 Brazil Network for Studies on Frailty in the Brazilian Elderly	773 (64%) 71.9 ± 5.9, *nr*	Continuous (0-60s)	Not stated	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional	4
Muir, 2010 Canada Project to Prevent Falls in Veterans	182 (30%) 79.9 ± 4.7, 60+	Binary (<10; 10s)	Open	Prospective (12 months)	Submitted monthly falls diary & received phone call every fall	0 falls; 1+ falls 0 or non-injurious falls; 1+ injurious falls	Longitudinal	6
Mulasso, 2017 Italy	192 (62%) 73 ± 6.2, 65+	Continuous (0-60s)	Not stated	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Longitudinal	6
Nevitt, 1989 USA	325 (81.8%) *nr* ± *nr,* 60+	Continuous (0-*nr*) Analysed binary (<2; ≥2s)	Not stated	Prospective (12 months)	Submitted weekly falls postcards and were contacted if missing postcard	0-1 falls; 2+ falls	Longitudinal	6
Niam, 1999 Singapore	68 (67.2%) 71.7 ± 8.1, 60-89	Continuous (0-60s)	Open	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional	1
Park, 2020 South Korea	39 (74.4%) 79 ± 5.3, 65-*nr*	Continuous (0-45s)	Open	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional	4

Porto, 2020 Brazil	101 (77.2%) 67.6 ± 5.0, 60-*nr*	Continuous (0-30s)	Not stated	Prospective (12 months)	Received monthly phone calls	0 falls; 1+ falls 1 fall; 2+ falls	Longitudinal Longitudinal	7
Rossat, 2011 France	7643 (50.5%) 70.9 ± 4.6, 65+	Binary (<5; 5s)	Not stated	Retrospective (12 months)	Self or proxy-reported questionnaire	0 falls; 1 fall; 2 falls; 3+ falls Continuous #of falls	Cross-sectional	4
Sampaio, 2013 Japan & Brazil	114 (80%) 71.8 ± 4.3, 65+	Continuous (0-30s)	Not stated	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	3
Shimada, 2009 Japan	455 (67.1%) 81.4 ±7.8, 65+	Continuous (0-120s) Analysed continuous and binary (≤3; >3s)	Not stated	Retrospective (12 months)	Self or proxy-reported questionnaire	0 falls; 1+ falls	Cross-sectional	4
Shimada, 2011 Japan	213 (61%) 80 ± 7.1, 65+	Continuous (0-120s)	Not stated	Retrospective (3 months)	Self or proxy-reported questionnaire	0 falls; 1+ falls	Cross-sectional	2
Shinohara, 2020 Japan	109 (84.4%) 76.9 ± 6.5, 65-*nr*	Continuous (0-30s)	Not stated	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	2
Shin, 2012 South Korea	356 (66.6%) 71.6 ± 4.9, 65+	Continuous (0-20s)	Open + Closed	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Cross-sectional	4
Swanenburg, 2013 Switzerland	146 (69.9%) 55 ± 22, 20-94	Continuous (0-*nr*) Analysed binary (<30; ≥30s)	Not stated	Retrospective (12 months)	Self-reported in interview	0-1 falls; 2+ falls	Cross-sectional	1
Thomas, 2005 United Kingdom	30 (53.9%) 80.4 ± 6.7, 65+	Continuous (0-*nr*) Analysed binary (≤1.02; >1.02)	Not stated	Retrospective (12 months)	Self-reported in interview; verified with medical notes and records	0-1 falls; 2+ falls	Cross-sectional	4
Tinetti, 1988 USA	336 (55%) 78.3 ± 5.1, 75+	Binary (unable to stand unsupported on one leg; able)	Open	Prospective (12 months)	Submitted bimonthly falls diaries & received bimonthly phone calls	0 falls; 1+ falls	Longitudinal	5

Yale Health and Aging Project								
Toulotte, 2006 France	40 (100%) 68.8 ± 5.6, 60+	Analysed # of times foot touched ground in 30s	Open + Closed	Retrospective (2 years)	Self-reported in two independent blinded interviews	0 falls; 1+ falls	Cross-sectional	4
Vellas, 1997 USA Albuquerque Falls Study	316 (59%) 72.7 ± 6.1, 60+	Binary (<5; 5s)	Open	Prospective (3 years)	Submitted bimonthly falls postcards & initiated phone call every fall	0 falls; 1+ falls 0 or non-injurious falls; 1+ injurious falls	Longitudinal	7
Vellas, 1998 USA Albuquerque Falls Study	405 (59%) 74 ± 6.7, 60+	Binary (<5; 5s)	Not stated	Prospective (2 years)	Submitted bimonthly falls postcards & initiated phone call every fall	0 falls; 1+ falls 0 or non-injurious falls; 1+ injurious falls	Longitudinal Longitudinal	6
Welmer, 2017 Sweden SNAC-K	2495 (61.9%) 72 ± 9.8, 60+	Continuous (0-60s)	Open	Prospective (3, 5, 10 years)	ICD-10 codes via linked health records	0 or non-injurious falls; 1+ injurious falls	Longitudinal	7
Yamada, 2010 Japan	171 (78.4%) 80.5 ± 5.6, 65+	Continuous (0-*nr*)	Open	Prospective (12 months)	Submitted monthly falls postcards	0 falls; 1+ falls	Longitudinal	5
Yamada, 2012 Japan	252 (76.6%) 78.3 ± 6.8, 65+	Continuous (0-60s)	Open	Retrospective (12 months)	Self-reported in interview	0 falls; 1+ falls	Cross-sectional Longitudinal	5
Yamada, 2020 Japan	471 (79.6%) 72.3 ± 7.3, 50-*nr*	Continuous (0-*nr*)	Open	Retrospective (12 months)	Self-reported questionnaire	0 falls; 1+ falls	Longitudinal	5

ICD International Classification of Diseases; **SNAC-K** Swedish National Study on Ageing and Care in Kungsholmen; **USA** United States of America
[a] Cross-sectional refers to analysis of balance and falls measures assessed at the same time. Longitudinal refers to analysis, where balance is assessed at baseline and falls are assessed after a given follow-up period.
For further study details, see Supplementary Table 1.

Table 2. Median balance time (seconds) by fallers and non-fallers

Author		**Sample size**	**Balance time in fallers (s) Median (Q1, Q3)**	**Balance time in non-fallers (s) Median (Q1, Q3)**	**P-value**
Shimada et al. (2011)		213	3 (IQR=4.0)	4 (IQR=6.0)	0.31
Sampaio et al. (2013)	Japan	40	15.2 (6.1, 29.0)	24.1 (9.2, 30.0)	0.56
	Brazil	74	13.9 (3.9-23.3)	12.7 (6.5, 26.2)	0.54
Ansai et al. (2020)		67	2.3 (1.4, 6.7)	3.1 (1.1, 9.4)	0.53
Heitmann et al. (1989)		110	4.62 (*nr*)	4.24 (*nr*)	>0.05
Niam & Wee (1999)		68	*nr*	*nr*	>0.05
Eto & Miyauchi (2018)		159	*nr*	*nr*	0.10
Hashidate et al. (2011)		30	*nr*	*nr*	>0.05
Arai et al. (2020) [a]		399	*nr*	*nr*	0.12
POOLED MEDIAN DIFFERENCE: 1.0 (-1.2, 8.9) [b]					

Q1= 25th percentile; Q3: 75th percentile; nr=not reported; IQR: interquartile range (Q3 – Q1)

[a] *Longitudinal study; all other studies are cross-sectional*

[b] *A positive difference indicates that non-fallers have longer balance time than fallers, while a negative difference indicates that fallers have longer balance times than non-fallers*

Table 3. Prognostic accuracy of balance test in predicting falls

Author	Cut-point	Sensitivity (%) [a]	Specificity (%) [b]	Positive Predictive Value (%) [c]	Negative Predictive Value (%) [d]	AUC (95% CI) [e]
ANY FALL (cross-sectional)						
Depasquale & Toscano (2009)	<6.5	48.3	89.7	82.4%	63.4%	0.766
Eto & Miyauchi (2018)	<55.4	-	-	-	-	0.533
Lin et al. (2004)	-	-	-	-	-	0.640
Rossat et al. (2010)	<5s	20.8 [f]	88.1 [f]	29.8 [f]	82.1 [f]	-
Shimada et al. (2009)	<3	51	61	-	-	-
ANY FALL (longitudinal)						
Beauchet et al. (2010)						
dominant	<5s	34.5 [f]	73.0 [f]	37.5 [f]	70.3 [f]	-
Bongue et al. (2011)						
non-dominant	<7.6s	46.0 [f]	65.3 [f]	38.4 [f]	72.0 [f]	0.56 (0.53,0.59)
dominant	<12.7s	60.9 [f]	49.1 [f]	36.0 [f]	72.7 [f]	0.55 (0.53,0.58)
Beauchet et al. (2010) [g] Bongue et al. (2011) [g]	Moved arms	50.6 [f]	59.7 [f]	37.2 [f]	72.0 [f]	-
Buatois et al. (2006)	<5s	28.1 [f]	65.2 [f]	25.8 [f]	67.7 [f]	-
Crenshaw et al. (2020)	per 1SD	-	-	-	-	0.56
Gerdhem et al. (2005)	per 1s	-	-	-	-	0.55 (0.51-0.60)
Lin et al. (2004)	per 1s	-	-	-	-	0.527
Muir et al. (2010)	<10s	74.4 [f]	46.2 [f]	50.9 [f]	70.6 [f]	-
Tinetti et al. (1988)	Unable	56.5 [f]	61.4 [f]	40.9 [f]	90.4 [f]	-
RECURRENT FALLS (cross-sectional)						
Jalali et al. (2015) [h]	<12.7	83.5	63	47.6	90.4	-
Swanenburg et al. (2013)	<30	61.1 [f]	52.5 [f]	28.2 [f]	81.6 [f]	-
Rossat et al. (2010)	<5s	26.6 [f]	87.2 [f]	12.1 [f]	94.7 [f]	-
Thomas & Lane (2005)	<1.02	67 (39-86)	89 (67-97)	-	-	-
RECURRENT FALLS (longitudinal)						
Beauchet et al. (2010)	<5	33	71.2	14.3	88.1	-
	Moved arms	55.9	58.2	16.2	90.1	-
Buatois et al. (2006)	<5s	16.7 [f]	90.3 [f]	15.4 [f]	91.1 [f]	-
Buatois et al. (2010)	<5s	42.1 [f]	68.2 [f]	12.9 [f]	91.3 [f]	-
Gerdhem et al. (2005)	<5s	28.8 [f]	89.1 [f]	14.5 [f]	95.1 [f]	-
INJURIOUS FALLS (longitudinal)						
Vellas et al. (1997)	<5	-	36	76	31	-

[a] Proportion of fallers who had a positive screening test (e.g. balance time < cut-point) (Parikh et al., 2008)
[b] Proportion of non-fallers who had a negative screening test (e.g. balance time ≥ cut-point) (Parikh et al., 2008)
[c] Proportion of those with a positive screening test (e.g. balance time < cut-point) who have a fall (Parikh et al., 2008)
[d] Proportion of those with a negative screening test (e.g. balance time ≥ cut-point) who do not have a fall (Parikh et al., 2008)
[e] Area under the curve
[f] Calculated using available data from paper
[g] Same sample; values for moved arms identical in both papers
[h] Values calculated from sample size; sensitivity, specificity, positive predictive value, and negative predictive value are incorrect in paper

3